

Schreiber, David

122982

**From:** Davis, Minh-Tam  
**Sent:** Thursday, May 20, 2004 12:48 PM  
**To:** Schreiber, David  
**Subject:** FW: Rush search request for 10/048046

A more complete request, with Room number and phone.  
Sorry.  
Tam

7/1999

-----Original Message-----

**From:** Davis, Minh-Tam  
**Sent:** Thursday, May 20, 2004 12:46 PM  
**To:** Schreiber, David  
**Subject:** Rush search request for 10/048046

Please perform a ~~Score/length~~ sequence for complement of SEQ ID NO:1, residues that encode amino acids 31-103, or amino acids 303-346, or amino acids 476-641 of SEQ ID NO:2.  
No specific minimum hit length is required. Maximum hit length 50  
~~Score over length value of 60% or greater.~~

Thank you.  
MINH TAM DAVIS  
ART UNIT 1642,  
ROOM 3A24, MB 3C18  
272-0830

100



QY	381	AlaArgAsnLysIleThrGlnAspMetLeuGlnProIysValArgSerPheSerAsp	400
Db	1231	GCAGGAATAAATCACTCAAGACATGCTGCAGCCCAAGATCAGGCGGTCTTTTCTGAT	1290
QY	401	GluGluGlySerSerGluAspLeuLeuGluLeuSerAspValAspSerGluSerSerAsp	420
Db	1291	GAAGAAGGGAGTTGAGAGGACCTGCTGGAGCTGTGACAGCTTGACAGTGTGCTCAGAC	1350
QY	421	IleSerGlnProTyrValValCysArgGlnCysProGluTyrArgArgGlnAlaGln	440
Db	1351	ATTAGCCAGCCATACGTGCTGTCGGGACGTGCTCTGAGTACAGAAGGCAGGGCGGCGCAG	1410
QY	441	ProProHisCysProAlaProGluGlyGluProGlyAlaProGlnAlaLeuGlyAspAla	460
Db	1411	CCTCCCACTGCCAGCACCCGAGGCGAGCCAGGAGCCACAGGCCCTTGGGGATGCA	1470
QY	461	ProSerThrSerValSerLeuThrThrAlaValGlnAspTyrValCysProLeuGlnGly	480
Db	1471	CCCTCCACGTCGCTCAGCCTGACGACAGCAGTCCAGGATTACGTGTGCCCTCTGCAAGGA	1530
QY	481	SerHisAlaLeuCysThrCysCysPheGlnProMetProAspArgAlaGluArgGlu	500
Db	1531	AGCCACGCCCTGTGCACCTGTGCTTCCAGCCCATGCCACCCGAGAGCGGAGCGGAG	1590
QY	501	GlnAspProArgValAlaPrGlnGlnCysAlaValCysLeuGlnProPheCysHisLeu	520
Db	1591	CAGACCCCGGTGTCGCCCTCAGCAGTGTGGGTGCTGCTGCGAGCCTTTCTGCCACCTG	1650
QY	521	TyrTrpGlyCysThrA-gthrGlyCyetyrGlyCysLeuAlaProPheCysGluLeuAsn	540
Db	1651	TACTGGGCTGCACCCGACCGGCTGCTACGGCTGCTGCGTCCCTGGCCCGCTTTTGTGAGCTCAAC	1710
QY	541	LeuGlyAspLysCysLeuAspGlyValLeuAsnAsnSerTyrGluSerAspIleLeu	560
Db	1711	CTGGGTGACAAAGTGTCTGGACGGGTGCTGAACCAACACAGCTACGAGTCAGACATCCTG	1770
QY	561	LysAsnTyrLeuAlaThrArgGlyLeuThrTrpLysAsnMetLeuThrGluSerLeuVal	580
Db	1771	AAGAATTACCTGGCAACCAAGAGTTTGACATGGAAAAACATGTTGACCGAGAGCCTCGTG	1830
QY	581	AlaLeuGlnArgGlyValPheLeuLeuSerAspTyrArgValThrGlyAspThrValLeu	600
Db	1831	GCTCTCCAGCGGGAGTGTCTGCTGCTGCTGATTACAGAGTACCGGAGACACCGTTCTG	1890
QY	601	CysTyrCysCysGlyLeuArgSerPheArgGluLeuThrTyrGlnTyrArgGlnAsnIle	620
Db	1891	TGTTACTGCTGTGCGCTCGCAGCTTCGCTGAGCTGACCTATCAGTATCGGCAGAACATT	1950
QY	621	ProAlaSerGluLeuProValAlaValThrSerArgProAspCysTyrTrpGlyArgAsn	640
Db	1951	CCTGCTTCCGAGTTGCCAGTGGCGGTAACTCCCGTCTCTGACTCTGCGGCCGCTAAC	2010
QY	641	CysArgThrGlnValLysAlaHisAlaMetLysPheAsnHisIleCysGluGlnThr	660
Db	2011	TGCCCATCTCAGGTGAAGCTCACACGCCATGAATTCATCATCATCTGTGTACAGACA	2070
QY	661	ArgPheLysAsn	684
Db	2071	AGGTTCAAAAC	2082

Search completed: May 20, 2004, 13:18:43  
Job time : 5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:10:28 ; Search time 3707.14 Seconds  
(without alignments)  
5822.506 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_1516\_2013  
Perfect score: 498  
Sequence: 1 tgcctctgcaggaagcca.....gctactgggcccgtactgc 498

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_scs.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_scs.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rtd.\*  
36: em\_htg\_nam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	21	4.2	30	6	I13395	I13395 Sequence 5
2	21	4.2	30	6	I13775	I13775 Sequence 5
3	21	4.2	30	6	I14876	I14876 Sequence 3
4	21	4.2	30	6	I68745	I68745 Sequence 5
5	21	4.2	30	6	I92686	I92686 Sequence 3
6	21	4.2	30	6	AR437164	AR437164 Sequence
7	20.8	4.2	50	6	AX165879	AX165879 Sequence
8	20.4	4.1	47	6	BD169427	BD169427 Genomes P
9	20.2	4.1	50	6	AR403372	AR403372 Sequence
10	20.2	4.1	50	6	BD068872	BD068872 Enzymatic
11	19.8	4.0	40	6	AX456252	AX456252 Sequence
12	19.8	4.0	42	6	AR078407	AR078407 Sequence
13	19.8	4.0	50	6	AR403397	AR403397 Sequence
14	19.8	4.0	50	6	BD068897	BD068897 Enzymatic
15	19.6	3.9	48	6	AR277861	AR277861 Sequence
16	19.4	3.9	42	6	E54530	E54530 Herpesvirus
17	19	3.8	36	6	AR206156	AR206156 Sequence
18	19	3.8	42	6	AR116992	AR116992 Sequence
19	19	3.8	42	6	AR116993	AR116993 Sequence
20	19	3.8	44	9	HUMSPARC01	M25737 Human osteo
21	18.8	3.8	50	6	I18452	I18452 Sequence 75
22	18.6	3.7	36	9	S81380	S81380 T cell anti
23	18.6	3.7	40	6	BD180755	BD180755 Array of
24	18.6	3.7	42	6	AX411814	AX411814 Sequence
25	18.6	3.7	47	6	AX004825	AX004825 Sequence
26	18.6	3.7	47	6	BD074255	BD074255 Method fo
27	18.4	3.7	36	6	BD271168	BD271168 Virus vec
28	18.4	3.7	36	6	BD271169	BD271169 Virus vec
29	18.4	3.7	36	6	AR264600	AR264600 Sequence
30	18.4	3.7	36	6	AR264601	AR264601 Sequence
31	18.4	3.7	38	6	AR057811	AR057811 Sequence
32	18.4	3.7	38	6	AR057818	AR057818 Sequence
33	18.4	3.7	38	6	AR057975	AR057975 Sequence
34	18.4	3.7	38	6	AR115569	AR115569 Sequence
35	18.4	3.7	38	6	AR115576	AR115576 Sequence
36	18.4	3.7	38	6	AR115733	AR115733 Sequence
37	18.4	3.7	38	6	AX634870	AX634870 Sequence
38	18.4	3.7	38	6	AX634877	AX634877 Sequence
39	18.4	3.7	38	6	AX635034	AX635034 Sequence
40	18.4	3.7	41	6	AX518482	AX518482 Sequence
41	18.4	3.7	45	6	AR282656	AR282656 Sequence
42	18.4	3.7	45	6	AR282657	AR282657 Sequence
43	18.4	3.7	45	6	AX370679	AX370679 Sequence
44	18.4	3.7	45	6	AX370680	AX370680 Sequence
45	18.4	3.7	47	6	AR289352	AR289352 Sequence

# ALIGNMENTS

RESULT 1  
I13395  
LOCUS I13395 30 bp DNA linear PAT 26-JUL-1995  
DEFINITION Sequence 5 from patent US 5436153.  
ACCESSION I13395  
VERSION I13395.1 GI:910736  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 30)  
AUTHORS Sprecher,C.A., Foster,D.C. and Norris,K.E.  
TITLE Human amyloid protein precursor homolog and Kunitz-type inhibitor  
JOURNAL Patent: US 5436153-A 5 25-JUL-1995;  
FEATURES Location/Qualifiers



source		1. .30	/organism="unknown"	/mol_type="unassigned DNA"	
ORIGIN					
Query Match		4.2%;	Score 21; DB 6; Length 30;		
Best Local Similarity		82.8%;	Pred. No. 1.5e+06;		
Matches		24; Conservative	0; Mismatches 5; Indels	0; Gaps	0;
QY		377 GTTACTGCTGTGGCTCGCAGCTTCCGT	405		
Db		1 GTTGTGCTGTGGCTCGCAGCTTCCGT	29		
RESULT 2					
LOCUS		I13775	30 bp	DNA	linear
DEFINITION		Sequence 5 from patent US 5441931.			
ACCESSION		I13775			
VERSION		I13775.1 GI:996205			
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unknown.			
REFERENCE		1 (bases 1 to 30)			
AUTHORS		Sprecher,C.A., Foster,D.C. and Norris,K.E.			
TITLE		Human amyloid protein precursor homologue and Kunitz-type inhibitors			
JOURNAL		Patent: US 5441931-A 5 15-AUG-1995;			
FEATURES		Location/Qualifiers			
source		1. .30			
/organism="unknown"					
/mol_type="unassigned DNA"					
ORIGIN					
Query Match		4.2%;	Score 21; DB 6; Length 30;		
Best Local Similarity		82.8%;	Pred. No. 1.5e+06;		
Matches		24; Conservative	0; Mismatches 5; Indels	0; Gaps	0;
QY		377 GTTACTGCTGTGGCTCGCAGCTTCCGT	405		
Db		1 GTTGTGCTGTGGCTCGCAGCTTCCGT	29		
RESULT 3					
LOCUS		I14876	30 bp	DNA	linear
DEFINITION		Sequence 3 from patent US 5455338.			
ACCESSION		I14876			
VERSION		I14876.1 GI:1249785			
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unknown.			
REFERENCE		1 (bases 1 to 30)			
AUTHORS		Sprecher,C.A., Kisiel,W. and Foster,D.C.			
TITLE		DNA encoding novel human kunitz-type inhibitors and methods relating thereto			
JOURNAL		Patent: US 5455338-A 3 03-OCT-1995;			
FEATURES		Location/Qualifiers			
source		1. .30			
/organism="unknown"					
/mol_type="unassigned DNA"					
ORIGIN					
Query Match		4.2%;	Score 21; DB 6; Length 30;		
Best Local Similarity		82.8%;	Pred. No. 1.5e+06;		
Matches		24; Conservative	0; Mismatches 5; Indels	0; Gaps	0;
QY		377 GTTACTGCTGTGGCTCGCAGCTTCCGT	405		
Db		1 GTTGTGCTGTGGCTCGCAGCTTCCGT	29		

source		1. .30	/organism="unknown"	/mol_type="unassigned DNA"	
ORIGIN					
Query Match		4.2%;	Score 21; DB 6; Length 30;		
Best Local Similarity		82.8%;	Pred. No. 1.5e+06;		
Matches		24; Conservative	0; Mismatches 5; Indels	0; Gaps	0;
QY		377 GTTACTGCTGTGGCTCGCAGCTTCCGT	405		
Db		1 GTTGTGCTGTGGCTCGCAGCTTCCGT	29		
RESULT 4					
LOCUS		I68745	30 bp	DNA	linear
DEFINITION		Sequence 5 from patent US 5677146.			
ACCESSION		I68745			
VERSION		I68745.1 GI:2830867			
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unknown.			
REFERENCE		1 (bases 1 to 30)			
AUTHORS		Sprecher,C.A., Foster,D.C. and Norris,K.E.			
TITLE		Human amyloid protein precursor homologue and kunitz-type inhibitor			
JOURNAL		Patent: US 5677146-A 5 14-OCT-1997;			
FEATURES		Location/Qualifiers			
source		1. .30			
/organism="unknown"					
/mol_type="unassigned DNA"					
ORIGIN					
Query Match		4.2%;	Score 21; DB 6; Length 30;		
Best Local Similarity		82.8%;	Pred. No. 1.5e+06;		
Matches		24; Conservative	0; Mismatches 5; Indels	0; Gaps	0;
QY		377 GTTACTGCTGTGGCTCGCAGCTTCCGT	405		
Db		1 GTTGTGCTGTGGCTCGCAGCTTCCGT	29		
RESULT 5					
LOCUS		I92686	30 bp	DNA	linear
DEFINITION		Sequence 3 from patent US 5728674.			
ACCESSION		I92686			
VERSION		I92686.1 GI:3937156			
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unknown.			
REFERENCE		1 (bases 1 to 30)			
AUTHORS		Sprecher,C.A., Kisiel,W. and Foster,D.C.			
TITLE		Inhibition of blood coagulation by human-kunitz-type inhibitors			
JOURNAL		Patent: US 5728674-A 3 17-MAR-1998;			
FEATURES		Location/Qualifiers			
source		1. .30			
/organism="unknown"					
/mol_type="unassigned DNA"					
ORIGIN					
Query Match		4.2%;	Score 21; DB 6; Length 30;		
Best Local Similarity		82.8%;	Pred. No. 1.5e+06;		
Matches		24; Conservative	0; Mismatches 5; Indels	0; Gaps	0;
QY		377 GTTACTGCTGTGGCTCGCAGCTTCCGT	405		
Db		1 GTTGTGCTGTGGCTCGCAGCTTCCGT	29		
RESULT 6					
LOCUS		AR437164	30 bp	DNA	linear
DEFINITION		Sequence 3 from patent US 6656746.			
ACCESSION		AR437164			
VERSION		AR437164.1 GI:40200261			
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unknown.			
REFERENCE		1 (bases 1 to 30)			
AUTHORS		Sprecher,C.A., Kisiel,W. and Foster,D.C.			
TITLE		Antibodies to human kunitz-type inhibitor			
JOURNAL		Patent: US 6656746-A 3 02-DEC-2003;			

FEATURES	Location/Qualifiers	
source	1..30	
	/organism="unknown"	
	/mol_type="genomic DNA"	
ORIGIN		
Query Match	4.2%; Score 21; DB 6; Length 30;	
Best Local Similarity	82.8%; Pred. No. 1.5e+06;	
Matches	24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	377 GTTACTGTGTGGCTGCGAGCTTCCT 405	
Db		
	1 GTTGTGTGTGCTGCTCGCAGCTTCCT 29	
RESULT 7		
AXI165879	50 bp DNA linear PAT 22-JUN-2001	
LOCUS	AXI165879	
DEFINITION	Sequence 1074 from Patent WO0138586.	
ACCESSION	AXI165879	
VERSION	AXI165879.1 GI:14546708	
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Fukariyota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 Shimkets, R.A. and Leach, M.	
JOURNAL	Nucleic acids containing single nucleotide polymorphisms and	
	Methods of use thereof	
	Patent: WO 0138586-A 1074 31-MAY-2001;	
JOURNAL	Curagen Corporation (US)	
FEATURES	Location/Qualifiers	
source	1..50	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
misc_feature	25..26	
	/note="Nucleotide deleted between bases 25 and 26	
variation	26	
	/note="single nucleotide polymorphism"	
ORIGIN		
Query Match	4.2%; Score 20.8; DB 6; Length 50;	
Best Local Similarity	70.0%; Pred. No. 1.7e+06;	
Matches	28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;	
QY	454 GTGGCCGTAAACATCCGCTCTGACTGCTACTGGGCCGTA 493	
Db		
	3 GTGGCCCTACACCCAGTCTCTCTGCGGAGGCTGTA 42	
RESULT 8		
BD169427	47 bp DNA linear PAT 17-JAN-2003	
LOCUS	BD169427	
DEFINITION	Genomes participating in rheumatoid arthritis, method of diagnosing	
	the same, method of judging the onset risk thereof, kit for	
	detecting and diagnosing the same, method of treating rheumatoid	
	arthritis and remedies therefor.	
ACCESSION	BD169427	
VERSION	BD169427.1 GI:27875239	
KEYWORDS	WO 0234912-A/4.	
SOURCE	synthetic construct	
ORGANISM	artificial sequences.	
REFERENCE	1 (bases 1 to 47)	
AUTHORS	Shiozawa, S. and Konishi, Y.	
TITLE	Genomes participating in rheumatoid arthritis, method of diagnosing	
	the same, method of judging the onset risk thereof, kit for	
	detecting and diagnosing the same, method of treating rheumatoid	
	arthritis and remedies therefor	
JOURNAL	Patent: WO 0234912-A 4 02-MAY-2002;	
COMMENT		
OS	Artificial Sequence	
PN	WO 0234912-A/4	
PD	02-MAY-2002	
PP	24-OCT-2001 WO 2001JP009313	
PR	24-OCT-2000 JP 00P 324296, 27-MAR-2001 JP 01P 090546 PR	
PI	SHUNICHI SHIOZAWA, YOSHITAKE KONISHI	
PC	C12N15/12, C07K14/47, C12Q1/68, G01N33/50, A61K38/17, A61K48/00 CC	
	Synthesized oligonucleotide	
FH	Key Location/Qualifiers	
FT	source 1..47	
	/organism="Artificial Sequence".	
FEATURES	Location/Qualifiers	
source	1..47	
	/organism="synthetic construct"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32630"	
ORIGIN		
Query Match	4.1%; Score 20.4; DB 6; Length 47;	
Best Local Similarity	65.2%; Pred. No. 2.2e+06;	
Matches	30; Conservative 0; Mismatches 16; Indels 0; Gaps 0;	
QY	27 GTGCACCTGTGCTTCCAGCCCATGCGCCGACCGGAGCGGAGCGC 72	
Db		
	2 GGTACATCCGCTTCTGCCCCAGCCAGCTGTTGTGGAGTGC 47	
RESULT 9		
AR403372	50 bp DNA linear PAT 18-DEC-2003	
LOCUS	AR403372	
DEFINITION	Sequence 1712 from patent US 6623962.	
ACCESSION	AR403372	
VERSION	AR403372.1 GI:40150822	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 50)	
AUTHORS	Akhtar, S., Fell, P. and McSwiggen, J.A.	
TITLE	Enzymatic nucleic acid treatment of diseases or conditions related	
	to levels of epidermal growth factor receptors	
JOURNAL	Patent: US 6623962-A 1712 23-SEP-2003;	
FEATURES	Location/Qualifiers	
source	1..50	
	/organism="unknown"	
	/mol_type="genomic DNA"	
ORIGIN		
Query Match	4.1%; Score 20.2; DB 6; Length 50;	
Best Local Similarity	75.8%; Pred. No. 2.4e+06;	
Matches	25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	268 GCACACGAGGTTTGACATGGAAAAACATGTTG 300	
Db		
	1 GCACCCAGAGTTTACCAGAGAAACACACGTTG 33	
RESULT 10		
BD068872	50 bp RNA linear PAT 27-AUG-2002	
LOCUS	BD068872	
DEFINITION	Enzymatic nucleic acid treatment of diseases or conditions related	
	to levels of epidermal growth factor receptors.	
ACCESSION	BD068872	
VERSION	BD068872.1 GI:22614475	
KEYWORDS	JP 2001511003-A/1712.	
SOURCE	unidentified	
ORGANISM	unclassified.	
REFERENCE	1 (bases 1 to 50)	
AUTHORS	Akhtar, S., Fell, P. and McSwiggen, J.A.	
TITLE	Enzymatic nucleic acid treatment of diseases or conditions related	

JOURNAL  
Patent: JP 2001511003-A 1712 07-AUG-2001;  
RIBOZYME PHARMACEUTICALS INC,ASTON UNIV  
OS Unidentified  
PN JP 2001511003-A/1712  
PD 07-AUG-2001  
PF 14-JAN-1998 JP 1998532913  
PR 31-JAN-1997 US 60/036476,04-DEC-1997 US 08/995162 PI  
SAGHIR AKHTAR,PATRICIA FELL,JAMES A MCSWIGGEN PC  
C12N9/00,C07K14/71  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Enzymatic nucleic acid treatment of diseases or conditions CC related to

CC levels of epidermal growth factor receptors  
FH Key Location/Qualifiers  
FT source 1..50  
FT /organism="Unidentified".  
FT Location/Qualifiers  
1..50  
/organism="unidentified"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:32644"

FEATURES  
source

ORIGIN

Query Match 4.1%; Score 20.2; DB 6; Length 50;  
Best Local Similarity 75.8%; Pred. No. 2.4e+06;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 268 GCACACGAGGTTTCACATGGAAACATGTTG 300  
|||||  
Db 1 GCACCCAGAGTTTACACAGAGAACACACGTTG 33  
|||||

RESULT 11  
AX456252/c  
LOCUS  
DEFINITION Sequence 110 from Patent WO0216944.  
ACCESSION AX456252  
VERSION AX456252.1 GI:21715166  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
AUTHORS Wood,K.V., Wood,M.G., Zhuang,Y. and Paguio,A.  
TITLE Synthetic nucleic acid molecule compositions and methods of preparation  
JOURNAL Patent: WO 0216944-A 110 28-FEB-2002;  
PROMEGA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..40  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="An oligonucleotide"

ORIGIN

Query Match 4.0%; Score 19.8; DB 6; Length 40;  
Best Local Similarity 69.2%; Pred. No. 3e+06;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 372 TCTGTGTTACTGCTGGCGCTCGCAGCTTCGCTGAGCT 410  
Db 39 TTTGTTGACAGTAGTATGATCTGACGAGCTTGGCTGAGCT 1  
|||||

RESULT 12  
AR078407  
LOCUS  
DEFINITION Sequence 26 from patent US 5962636.  
ACCESSION AR078407  
VERSION AR078407.1 GI:10005153

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE  
1 (bases 1 to 42)  
AUTHORS Bachmaier,K., Hessel,A.John., Neu,N. and Penninger,J.Martin.  
TITLE Peptides capable of modulating inflammatory heart disease  
JOURNAL Patent: US 5962636-A 26 05-OCT-1999;  
FEATURES Location/Qualifiers  
source 1..42  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 4.0%; Score 19.8; DB 6; Length 42;  
Best Local Similarity 77.4%; Pred. No. 3e+06;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 115 CTGCAGCCTTCTGCCACCTGTACTGGGCT 145  
|||||  
Db 12 CAGCACCTGTTTGCACACTATGCTGGGCT 42  
|||||

RESULT 13  
AR403397  
LOCUS  
DEFINITION Sequence 1737 from patent US 6623962.  
ACCESSION AR403397  
VERSION AR403397.1 GI:40150847  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE  
1 (bases 1 to 50)  
AUTHORS Akhtar,S., Fell,P. and McSwiggen,J.A.  
TITLE Enzymatic nucleic acid treatment of diseases or conditions related to levels of epidermal growth factor receptors  
JOURNAL Patent: US 6623962-A 1737 23-SEP-2003;  
FEATURES Location/Qualifiers  
source 1..50  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 4.0%; Score 19.8; DB 6; Length 50;  
Best Local Similarity 63.8%; Pred. No. 3e+06;  
Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 224 TGAACACACAGCTACGAGTCAGACATCTCGAAGATTACCTGGCA 270  
|||||  
Db 4 TGAAGAAGAACACACAGAGAACACACGTTGTGTGTACATTACCTGTA 50  
|||||

RESULT 14  
BD068897  
LOCUS  
DEFINITION Enzymatic nucleic acid treatment of diseases or conditions related to levels of epidermal growth factor receptors.  
ACCESSION BD068897  
VERSION BD068897.1 GI:22614500  
KEYWORDS JP 2001511003-A/1737.  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

REFERENCE  
1 (bases 1 to 50)  
AUTHORS Akhtar,S., Fell,P. and McSwiggen,J.A.  
TITLE Enzymatic nucleic acid treatment of diseases or conditions related to levels of epidermal growth factor receptors  
JOURNAL Patent: JP 2001511003-A 1737 07-AUG-2001;  
RIBOZYME PHARMACEUTICALS INC,ASTON UNIV  
COMMENT OS Unidentified  
PN JP 2001511003-A/1737  
PD 07-AUG-2001

PF 14-JAN-1998 JP 1998532913  
PR 31-JAN-1997 US 60/036476,04-DEC-1997 US 09/985162 P1  
SAGHIR AKHTAR,PATRICIA FELL,JAMES A MCSWIGGEN PC  
C12N9/00,C07K14/71

CC Strandedness: Single;  
CC Topology: Linear;  
CC Enzymatic nucleic acid treatment of diseases or conditions CC  
related to

CC levels of epidermal growth factor receptors  
FH Key Location/Qualifiers  
FT source 1..50  
/organism='Unidentified'.  
Location/Qualifiers

1..50  
/organism="unidentified"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 4.0%; Score 19.8; DB 6; Length 50;  
Best Local Similarity 63.8%; Pred. No. 3e+06;  
Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 224 TGACAAACAGCAGTCAGATCCTGGAAGAAATACCTGGCA 270

DB 4 TGAGAGAGACACAGAGAAACACACGTTGTGTACATTACCTGGTA 50

RESULT 15

AR277861/c  
LOCUS AR277861 48 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 20 from patent US 6511803.  
ACCESSION AR277861  
VERSION AR277861.1 GI:29711785

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 48)  
AUTHORS Church,G.M. and Mitra,R.D.  
TITLE Replica amplification of nucleic acid arrays  
JOURNAL Patent: US 6511803-A 20 28-JAN-2003;  
FEATURES Location/Qualifiers

source 1..48  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 3.9%; Score 19.6; DB 6; Length 48;  
Best Local Similarity 73.5%; Pred. No. 3.4e+06;  
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 36 CTGCTTCAGCCCATGCCCGACCGGAGAGCGGAG 69

DB 43 CTCCTTCAGCTTCGCCCGGAGGAGGAGCGGAG 10

Search completed: May 26, 2004, 07:39:20  
Job time : 3709.14 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:01:38 ; Search time 420.572 Seconds  
(without alignment)  
5030.293 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_1516\_2013

Perfect score: 498

Sequence: 1 tgcctctgcgaaggaagca.....Gctactggggccgtaactgc 498

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	21.6	4.3	31	2	AAX06374 Human bia
C 2	21.2	4.3	50	2	AAV04790 Primer 95
C 3	21.2	4.3	50	2	AAV31045 Expressio
C 4	21	4.2	30	2	AAG66249 Sequence
C 5	21	4.2	30	2	AAG90470 Human Kun
C 6	21	4.2	50	4	AAL32458 Human SNP
C 7	21	4.2	50	4	AAL32043 Human SNP
C 8	20.8	4.2	41	6	ABV74248 Human mac
C 9	20.8	4.2	48	7	ACC62681 Human ant
C 10	20.8	4.2	48	7	ACC62643 Human ant
C 11	20.8	4.2	50	5	ABL01083 Human SNP
C 12	20.4	4.1	47	6	AAI47188 Human the
C 13	20.2	4.1	50	2	AAV98932 Human RGF
C 14	20	4.0	50	6	ABZ00921 Human leu
C 15	19.8	4.0	40	6	ABI99142 Green/red
C 16	19.8	4.0	42	2	AAZ28188 Human bet
C 17	19.8	4.0	42	3	AAZ99171 Human Dep
C 18	19.8	4.0	50	2	AAV98957 Human RGF
C 19	19.8	4.0	50	4	AAL34581 Human SNP
C 20	19.8	4.0	50	4	AAL34023 Human syn
C 21	19.6	3.9	41	6	ABA05129 Human syn
C 22	19.6	3.9	48	3	AAZ99461 Primer Pu
C 23	19.6	3.9	48	4	AAI12549 PCR prime

24 19.6 3.9 50 6 ABZ07831  
c 25 19.2 3.9 41 6 ABV74249  
26 19.2 3.9 50 4 AAL28785  
c 27 19.2 3.9 50 4 AAL34331  
c 28 19.2 3.9 50 4 AAL34262  
29 19 3.8 50 4 AAL34327  
30 18.8 3.8 41 6 ABL55510  
31 18.8 3.8 48 7 ACC62644  
32 18.8 3.8 48 7 ACC62693  
33 18.8 3.8 48 7 ACC62655  
34 18.8 3.8 48 7 ACC62682  
35 18.6 3.7 42 6 AAD39247  
c 36 18.6 3.7 44 2 AAQ51167  
c 37 18.6 3.7 47 2 AAX56654  
38 18.6 3.7 50 6 ABX03168  
39 18.6 3.7 50 6 ABX03263  
40 18.6 3.7 50 6 ABZ02873  
41 18.4 3.7 36 3 AAD00854  
c 42 18.4 3.7 36 3 AAD00855  
c 43 18.4 3.7 38 2 AAT53816  
c 44 18.4 3.7 38 2 AAT53973  
c 45 18.4 3.7 38 2 AAT53809

## ALIGNMENTS

RESULT 1  
AAX06374/c  
ID AAX06374 standard; DNA; 31 BP.

XX AC AAX06374;

XX DT 31-MAR-1999 (first entry)

XX DE Human biallelic polymorphic DNA fragment SGC33319.

XX KW Polymorphism; biallelic; paternity testing; forensic; genetic mapping;  
XX KM phenotypic typing; medicament; disease; marker; human; ss.

XX OS Homo sapiens.

XX FN WO9858529-A2.

XX PD 30-DEC-1998.

XX PF 22-JUN-1998; 98WO-US012930.

XX PR 24-JUN-1997; 97US-0050594P.

XX PA (AFFY-) AFFYMETRIX INC.

XX PI Lipshutz RJ, Chee M, Fan J, Berno A;

XX DR WPI, 1999-080963/07.

XX PT New nucleic acid segments containing polymorphic sites - used for, e.g.  
XX PT detecting a disease phenotype, in forensics, paternity testing or genetic  
XX PT mapping of phenotypic traits.

XX PS Claim 1; Page 23; 61pp; English.

XX CC Sequences AAX06101-X06558 represent human DNA fragments which contain  
XX CC biallelic polymorphic markers. The base occupying the polymorphic site is  
XX CC indicated by the appropriate IUPAC-IUB ambiguity code. These fragments  
XX CC can be used in a method for determining polymorphic forms in an  
XX CC individual. The invention further provides computer-readable storage  
XX CC medium for storing data for access by an application programme being  
XX CC executed on a data processing system. Such a method comprises a data  
XX CC structure stored in the computer-readable storage medium, the data  
XX CC structure including information resident in a database used by the  
XX CC application programme and including records, each record comprising  
XX CC information identifying a polymorphism shown in the above sequences. The

CC products and methods can be used for analysing polymorphic sites in  
 CC individuals for testing for the presence of a disease phenotype or in  
 CC forensics, paternity testing or genetic mapping of phenotypic traits.  
 CC They can also be used for the production of polypeptides expressed by  
 CC variant genes and for the production of transgenic animals. The nucleic  
 CC acid segments can also be used in the manufacture of medicaments for the  
 CC treatment or prophylaxis of diseases

SQ Sequence 31 BP; 5 A; 8 C; 8 G; 9 T; 0 U; 1 Other;  
 Query Match 4.3%; Score 21.6; DB 2; Length 31;  
 Best Local Similarity 80.0%; Pred. No. 2.5e+04;  
 Matches 24; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 230 ACAACAGCTACGAGTCAGACATCTCTGAAGA 259  
 DB 30 ACATCAGCTACGAGCTCGAATCTCTGGAGA 1

RESULT 2  
 AAV04790/c  
 ID AAV04790 standard; cDNA; 50 BP.  
 AC AAV04790;  
 XX  
 DT 10-JUN-1998 (first entry)  
 XX  
 DE Primer 95-1414 for green fluorescent protein.  
 XX  
 KW PCR primer; green fluorescent protein; GFP; recombinant fungal host cell;  
 KW cryptic splice site; hormone production; enzyme production;  
 KW receptor production; reporter production; ss.  
 XX  
 OS Synthetic.  
 OS Aequorea victoria.  
 XX  
 PN WO9749821-A1.  
 XX  
 PD 31-DEC-1997.  
 XX  
 PF 20-JUN-1997; 97WO-US010630.  
 XX  
 PR 27-JUN-1996; 96US-0020930P.  
 XX  
 PA (NOVO ) NOVO NORDISK BIOTECH INC.  
 XX  
 PI Thompson SA;  
 XX  
 DR WPI; 1998-077184/07.  
 XX  
 PT Recombinant heterologous polypeptide expression in fungal host cells -  
 PT comprises modifying at least 1 cryptic splice site in nucleic acid  
 PT sequence encoding heterologous polypeptide, to provide correct  
 PT expression.  
 XX  
 PS Example 1; Page 22; 59pp; English.  
 XX  
 CC This sequence represents a primer used in the construction of a vector  
 CC containing the Aequorea victoria green fluorescent protein (GFP). The  
 CC vector can be used in the method of the invention for obtaining a novel  
 CC recombinant fungal host cell (FHC), which comprises introducing a nucleic  
 CC acid sequence (NAS) encoding a heterologous polypeptide, where at least 1  
 CC cryptic splice site (CSS) is modified, into a FHC. The products can be  
 CC used for the recombinant production of hormones, enzymes (e.g.  
 CC aminopeptidases, amylases, carbohydrases, carboxypeptidases, catalases,  
 CC cellulases, chitinases, cutinases, deoxyribonucleases, esterases, alpha-  
 CC galactosidases, beta-galactosidases, glucoamylases, alpha-glucosidases,  
 CC beta-glucosidases, haloperoxidases, invertases, lactases, lipases,  
 CC mannosidases, mutanases, oxidases, pectinolytic enzymes, peroxidases,  
 CC phytases, polyphenoloxidases, proteolytic enzymes, ribonucleases or  
 CC xylanases), receptors or reporters (e.g. Aequorea victoria GFP). The CSS  
 CC modification prevents improper splicing of precursor mRNA for  
 CC heterologous expression by FHC, therefore preventing the production of

CC biologically inactive polypeptides, or the production of several  
 CC populations of polypeptide products with varying biological activity

SQ Sequence 50 BP; 8 A; 13 C; 18 G; 11 T; 0 U; 0 Other;  
 Query Match 4.3%; Score 21.2; DB 2; Length 50;  
 Best Local Similarity 69.0%; Pred. No. 3.8e+04;  
 Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 396 CAGCTTCCGTCGAGCTGACCTATCAGTATCGGCAGAACATTCC 437  
 DB 50 CGGCAGGTTCACTGCTCGCGACCATTTATCAGCAGAACACTCC 9

RESULT 3  
 AAV31045/c  
 ID AAV31045 standard; cDNA; 50 BP.  
 XX  
 AC AAV31045;  
 XX  
 DT 17-AUG-1998 (first entry)  
 XX  
 DE Expression vector pShTh49 construction primer 95-1414.  
 XX  
 KW Fungal host cell; cryptic splice site; filamentous fungus; primer;  
 KW splicing; biologically active; construction; expression vector; ss.  
 XX  
 OS Synthetic.  
 OS Aequorea victoria.  
 XX  
 PN US5770371-A.  
 XX  
 PD 23-JUN-1998.  
 XX  
 PF 27-JUN-1996; 96US-00672158.  
 XX  
 PR 27-JUN-1996; 96US-00672158.  
 XX  
 PA (NOVO ) NOVO NORDISK BIOTECH INC.  
 XX  
 PI Thompson S;  
 XX  
 DR WPI; 1998-376791/32.  
 XX  
 PT DNA for transforming filamentous fungi - with cryptic splice sites  
 PT modified to prevent incorrect splicing in fungi.  
 XX  
 PS Example 1; Col 14; 27pp; English.  
 XX

CC The present sequence represents a primer used in the construction of the  
 CC expression vector pShTh49 in an example from the present invention. The  
 CC present invention describes a nucleic acid sequence with at least one  
 CC modified cryptic splice site, where the nucleic acid sequence encodes a  
 CC biologically active polypeptide and the cryptic splice site is modified  
 CC by (a) replacing at least one cryptic consensus sequence with a non-  
 CC consensus sequence; and/or (b) replacing a first region of a cryptic  
 CC intron with a second region that has a G+C content of 40-70%. The fungal  
 CC host cells may be used to produce biologically active polypeptides,  
 CC preferably hormones, enzymes, receptors or reporters. Problems of  
 CC incorrect splicing of precursor mRNA in filamentous fungal cells can be  
 CC overcome

SQ Sequence 50 BP; 8 A; 13 C; 18 G; 11 T; 0 U; 0 Other;  
 Query Match 4.3%; Score 21.2; DB 2; Length 50;  
 Best Local Similarity 69.0%; Pred. No. 3.8e+04;  
 Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 396 CAGCTTCCGTCGAGCTGACCTATCAGTATCGGCAGAACATTCC 437  
 DB 50 CGGCAGGTTCACTGCTCGCGACCATTTATCAGCAGAACACTCC 9

```
RESULT 4
AAQ66249
ID AAQ66249 standard; cDNA; 30 BP.
XX
AC AAQ66249;
XX
XX
DT 25-MAR-2003 (revised)
DT 29-DEC-1994 (first entry)
XX
XX
DE Sequence of probe ZC4792 for screening for a protease inhibitor from the
DE Kunitz family from human placenta.
XX
XX
KW Kunitz-type protease inhibitor; serine protease inhibitor; therapy;
KW amyloid precursor; probe; ss.
XX
OS Synthetic.
XX
PN WO9412637-A2.
XX
PD 09-JUN-1994.
XX
PF 01-DEC-1993; 93WO-US011696.
XX
PR 02-DEC-1992; 92US-00985692.
PR 19-NOV-1993; 93US-00155331.
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Sprecher CA, Foster DC, Norris K;
XX
XX WPI; 1994-200265/24.
XX
XX New human Kunitz type protease inhibitor and related DNA - for treating
PT pancreatitis and other disorders involving serine protease, also new
PT amyloid protein precursor homologues including the inhibitor in its
PT sequence.
XX
PS Example; Page 47; 70pp; English.
XX
XX A DNA segment encoding a human Kunitz-type inhibitor which comprises bps
CC 171-331 or 159-331 of AAQ66247 is claimed. Also claimed is the DNA
CC encoding AAR53776 from AA 56-110, from 53-110 or from 57-110 and which
CC further contains a Glu on the NH2 terminus. Isolated human Kunitz- type
CC inhibitor which comprises the aforementioned sequences is also claimed.
CC To obtain a cDNA sequence encoding a protease inhibitor from the Kunitz
CC family, a human placenta cDNA library in lambda gt11 was screened using
CC the radio-labelled probe ZC4792 (AAQ66249). Plasmids contg. two resultant
CC clones - ZGKI13 and ZGKI20 - were deposited on October 14th 1992 with the
CC ATCC under accession nos. 69090 and 69089 respectively. Plasmids ZGKI13
CC and ZGKI20 were shown to contain the identical partial sequence shown in
CC AAQ66247 and AAR53776. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 30 BP; 2 A; 10 C; 8 G; 10 T; 0 U; 0 Other;
Query Match 4.2%; Score 21; DB 2; Length 30;
Best Local Similarity 82.8%; Pred. No. 3.7e+04;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 377 GTTACTGCTGTCGCTCGCAGCTTCGGT 405
DB 1 GTTGTGCTGTTGCTTCGCCAGCTTCGGT 29
XX
XX
RESULT 5
AAQ90470
ID AAQ90470 standard; DNA; 30 BP.
XX
AC AAQ90470;
XX
XX
DT 21-JAN-1996 (first entry)
XX
XX Human Kunitz-type protease-inhibitor antisense DNA probe ZC4792.
DE
```

```
XX
KW Kunitz-type inhibitor; protease-inhibitor; enzyme-inhibitor;
KW anticoagulant; deep vein thrombosis; antisense DNA probe; ss.
XX
OS Synthetic.
XX
XX WO9512674-A1.
XX
XX 11-MAY-1995.
XX
XX 02-NOV-1994; 94WO-US012609.
XX
XX 05-NOV-1993; 93US-00147710.
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA (UYNE-) UNIV NEW MEXICO STATE.
XX
XX Sprecher CA, Kisiel W, Foster DC;
XX
XX WPI; 1995-193821/25.
XX
XX Human Kunitz-type protease inhibitors - used as anticoagulants and in the
PT treatment of deep vein thrombosis.
XX
XX Disclosure; Page 51; 65pp; English.
XX
XX Poly(A)+ RNAs from a variety of human tissues were screened using this
CC antisense 30-mer oligonucleotide as an initial step in the cloning of a
CC novel human Kunitz inhibitor cDNA
XX
SQ Sequence 30 BP; 2 A; 10 C; 8 G; 10 T; 0 U; 0 Other;
Query Match 4.2%; Score 21; DB 2; Length 30;
Best Local Similarity 82.8%; Pred. No. 3.7e+04;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 377 GTTACTGCTGTCGCTCGCAGCTTCGGT 405
DB 1 GTTGTGCTGTTGCTTCGCCAGCTTCGGT 29
XX
XX
RESULT 6
AAL32458
ID AAL32458 standard; DNA; 50 BP.
XX
AC AAL32458;
XX
XX 24-JAN-2002 (first entry)
XX
XX Human SNP oligonucleotide #5666.
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200147944-A2.
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US035498.
XX
XX 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
XX (CURA-) CURAGEN CORP.
XX
```

PI Shimkets RA, Leach M;  
XX WPI; 2001-465210/50.  
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
PT autoimmune diseases and infections.  
XX Claim 1; Page 3019; 4143pp; English.  
PS  
XX The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinases, cytokines, interferons, interleukins, G-  
CC protein coupled receptors and thioesterases. The present sequence is one  
CC such oligonucleotide. The oligonucleotides and the peptides encoded by  
CC associated with inappropriate expression of the proteins listed above.  
CC Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms  
XX  
SQ Sequence 50 BP; 12 A; 17 C; 10 G; 11 T; 0 U; 0 Other;  
Query Match 4.2%; Score 21; DB 4; Length 50;  
Best Local Similarity 73.0%; Pred. No. 4.3e+04;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 125 TCTGCCACCTGACTGAGGCTACACACGCGGTG 161  
DB 13 TCTGCCACCTGACTGAGGCTACACACGCGGTG 49  
RESULT 7  
ID AAL32043  
XX AAL32043 standard; DNA; 50 BP.  
AC AAL32043;  
XX  
DT 24-JAN-2002 (first entry)  
DE Human SNP oligonucleotide #5251.  
XX  
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinase; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-US035498.  
XX  
PR 28-DEC-1999; 99US-0173419P.  
PR 27-DEC-2000; 2000US-00173419.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX Shimkets RA, Leach M;  
PI WPI; 2001-465210/50.  
DR

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
PT autoimmune diseases and infections.  
XX Claim 1; Page 2898; 4143pp; English.  
PS  
XX The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinases, cytokines, interferons, interleukins, G-  
CC protein coupled receptors and thioesterases. The present sequence is one  
CC such oligonucleotide. The oligonucleotides and the peptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of the proteins listed above.  
CC Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms  
XX  
SQ Sequence 50 BP; 8 A; 15 C; 17 G; 10 T; 0 U; 0 Other;  
Query Match 4.2%; Score 21; DB 4; Length 50;  
Best Local Similarity 82.8%; Pred. No. 4.3e+04;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 24 CCTGTGACCTGCTGCTTCCAGCCCATGC 52  
DB 11 CATGGGACCTGCTGCTTCTAGCCACTGC 39  
RESULT 8  
ABV74248/C  
ID ABV74248 standard; DNA; 41 BP.  
XX  
AC ABV74248;  
XX  
DT 02-JAN-2003 (first entry)  
DE Human macro protein 0986-9.24 probe SEQ ID NO 8.  
XX  
DE Human macro protein 0986-9.24; tumour; haemopathy; development disorder;  
KW human immunodeficiency virus; HIV; infection immunological disease;  
KW inflammation; probe; ss.  
XX  
OS Homo sapiens.  
XX  
FN CN1352149-A.  
XX  
PD 05-JUN-2002.  
XX  
PF 10-NOV-2000; 2000CN-00127337.  
XX  
PR 10-NOV-2000; 2000CN-00127337.  
XX  
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.  
XX  
PI Mao Y, Xie Y;  
XX WPI; 2002-699451/76.  
XX  
DR New polypeptide-human macro protein 0986-9.24 and polynucleotide for  
PT encoding such polypeptide.  
XX  
PS Example 6; Page 19 (Disclosure); 32pp; Chinese.  
XX  
XX The invention relates to human macro protein 0986-9.24, polynucleotides  
CC encoding this polypeptide and DNA recombinant processes to produce the  
CC polypeptide. The present invention also discloses applying the



CC polypeptide in treating various diseases, such as malignant tumours,  
CC haenopathy, development disorder, human immunodeficiency virus (HIV)  
CC infection, immunological diseases and various inflammations. The present  
CC sequence is that of a human macro protein 0986-9.24 probe, useful in  
CC examples of the invention  
SQ Sequence 41 BP; 8 A; 9 C; 13 G; 11 T; 0 U; 0 Other;  
Query Match 4.2%; Score 20.8; DB 6; Length 41;  
Best Local Similarity 70.0%; Pred. No. 4.6e-04;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 420 GTATCGGAGACATTCCTGCTTCGAGTTGCCAGTGGCC 459  
Db 41 GCATGGGCAAAACATTCATCCATCCCTAGAGTTCACAGTGCC 2  
RESULT 9  
ACC62681  
ID ACC62681 standard; DNA; 48 BP.  
XX  
AC ACC62681;  
XX  
DT 30-JUN-2003 (first entry)  
XX  
DE Human antibody heavy chain framework 1 primer TMX24VH1b SEQ ID NO:64.  
XX  
KW Engineered template; single primer amplification; antibody library;  
KW nucleic acid amplification; primer; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO2003025202-A2.  
XX  
PD 27-MAR-2003.  
XX  
PF 19-SEP-2002; 2002WO-US029889.  
XX  
PR 19-SEP-2001; 2001US-0323455P.  
XX  
PA (ALEX-) ALEXION PHARM INC.  
XX  
PI Bowdish KS, Frederickson S, Maruyama T, Lin Y, Renshaw M;  
XX WPI; 2003-313359/30.  
XX  
PT Amplifying nucleic acid by contacting engineered nucleic acid strand  
PT having predetermined sequence at one end and sequence complementary to  
PT predetermined sequence at other end, with primer having predetermined  
PT sequence.  
PS Example 3; Page 29; 68pp; English.  
XX  
CC The present invention describes a method (M1) for amplifying a nucleic  
CC acid strand. M1 comprises providing an engineered nucleic acid strand (S)  
CC having a predetermined sequence at one end and a sequence complementary  
CC to the predetermined sequence at the other, and contacting (S) with a  
CC primer having the predetermined sequence in the presence of a polymers  
CC and nucleotides under conditions suitable for polymerisation of the  
CC nucleotides. Also described is an engineered nucleic acid strand (I)  
CC having a predetermined sequence at one end and a sequence complementary  
CC to the predetermined sequence at the other end. M1 is useful for  
CC amplifying a nucleic acid. M1 can be used for producing an antibody  
CC library. M1 is useful for preparing amplified products that can be  
CC ligated into a suitable expression vector, where the vector can be used  
CC to transform an appropriate host organism to produce the polypeptide or  
CC protein encoded by the target sequence. M1 is useful to amplify a family  
CC of related sequences to build a complex library such as, for example an  
CC antibody library. M1 is useful not only for producing large amounts of  
CC one target nucleic acid sequence, but also for amplifying simultaneously  
CC more than one different target nucleic acid sequence located on the same  
CC or different nucleic acid molecules. ACC62635 to ACC62753 and ABR54841 to

CC ABR54998 represent sequence used in the exemplification of the present  
CC invention  
XX Sequence 48 BP; 8 A; 10 C; 20 G; 10 T; 0 U; 0 Other;  
SQ  
Query Match 4.2%; Score 20.8; DB 7; Length 48;  
Best Local Similarity 70.0%; Pred. No. 4.8e+04;  
Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 50 TGCCCGACGGAGCGGAGCGGAGCGAGCAGGACCGCGGTGT 89  
Db 5 TGGCCGTTGAGAGAGAGAGTCTCGAGCAGGTCCAGCTTGT 44  
RESULT 10  
ACC62643  
ID ACC62643 standard; DNA; 48 BP.  
XX  
AC ACC62643;  
XX  
DT 30-JUN-2003 (first entry)  
XX  
DE Human antibody heavy chain VH framework 1 primer TMX24VH1b SEQ ID NO:26.  
XX  
KW Engineered template; single primer amplification; antibody library;  
KW nucleic acid amplification; primer; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO2003025202-A2.  
XX  
PD 27-MAR-2003.  
XX  
PF 19-SEP-2002; 2002WO-US029889.  
XX  
PR 19-SEP-2001; 2001US-0323455P.  
XX  
PA (ALEX-) ALEXION PHARM INC.  
XX  
PI Bowdish KS, Frederickson S, Maruyama T, Lin Y, Renshaw M;  
XX WPI; 2003-313359/30.  
XX  
PT Amplifying nucleic acid by contacting engineered nucleic acid strand  
PT having predetermined sequence at one end and sequence complementary to  
PT predetermined sequence at other end, with primer having predetermined  
PT sequence.  
PS Example 2; Page 23; 68pp; English.  
XX  
CC The present invention describes a method (M1) for amplifying a nucleic  
CC acid strand. M1 comprises providing an engineered nucleic acid strand (S)  
CC having a predetermined sequence at one end and a sequence complementary  
CC to the predetermined sequence at the other, and contacting (S) with a  
CC primer having the predetermined sequence in the presence of a polymers  
CC and nucleotides under conditions suitable for polymerisation of the  
CC nucleotides. Also described is an engineered nucleic acid strand (I)  
CC having a predetermined sequence at one end and a sequence complementary  
CC to the predetermined sequence at the other end. M1 is useful for  
CC amplifying a nucleic acid. M1 can be used for producing an antibody  
CC library. M1 is useful for preparing amplified products that can be  
CC ligated into a suitable expression vector, where the vector can be used  
CC to transform an appropriate host organism to produce the polypeptide or  
CC protein encoded by the target sequence. M1 is useful to amplify a family  
CC of related sequences to build a complex library such as, for example an  
CC antibody library. M1 is useful not only for producing large amounts of  
CC one target nucleic acid sequence, but also for amplifying simultaneously  
CC more than one different target nucleic acid sequence located on the same  
CC or different nucleic acid molecules. ACC62635 to ACC62753 and ABR54841 to  
CC ABR54998 represent sequence used in the exemplification of the present  
CC invention  
XX

SQ Sequence 48 BP; 8 A; 10 C; 20 G; 10 T; 0 U; 0 Other;  
 Query Match 4.2%; Score 20.8; DB 7; Length 48;  
 Best Local Similarity 70.0%; Pred. No. 4.9e+04;  
 Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 50 TGCCCGACCGAGAGCGGAGCGGAGCAGGACCCGCGTGT 89  
 |||||  
 Db 5 TGGCGCTTGGGAAGAGGAGTGCTCGAGCAGGTCCAGCTGT 44  
 |||||

RESULT 11  
 ABL01083  
 ID ABL01083 standard; DNA; 50 BP.  
 XX AC ABL01083;  
 XX AC  
 XX AC  
 DT 05-MAR-2002 (first entry)  
 XX DE  
 XX DE  
 XX Human SNP involving a gap oligonucleotide SEQ ID NO:1074.  
 XX Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;  
 XX immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;  
 KW autoimmune disease; inflammation; cancer; nervous system disease;  
 KW infection; polymorphic protein; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX W0200138586-A2.  
 PN  
 XX 31-MAY-2001.  
 PD  
 XX  
 XX 22-NOV-2000; 2000WO-US032311.  
 XX  
 XX 24-NOV-1999; 99US-0167383P.  
 PR  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Shimkets RA, Leach M;  
 PI  
 XX WPI; 2001-355949/37.  
 DR  
 XX  
 XX Isolated human nucleic acids comprising one or more single nucleotide  
 PT polymorphisms, useful for treating a subject suffering from a pathology,  
 PT e.g. autoimmune diseases, ascribed to the presence of a sequence  
 PT polymorphism.  
 PT  
 XX  
 PS Claim 1; Page 575; 674pp; English.  
 XX  
 CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides  
 CC comprising one or more single nucleotide polymorphisms (SNPs). ABL056531  
 CC to ABL056903 represent human peptides encoded by some of the SNP  
 CC oligonucleotides. The sequences from the present invention can have  
 CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and  
 CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides  
 CC and antibodies from the present invention can be used for treating a  
 CC subject suffering from, at risk for, or suspected of, suffering from a  
 CC pathology ascribed to the presence of a sequence polymorphism. The  
 CC pathology may be autoimmune diseases, inflammation, cancer, diseases of  
 CC the nervous system, and infection by pathogenic microorganisms. The SNPs  
 CC are also useful for determining which forms of a characterised  
 CC polymorphism are present in individuals. The antibodies may be used in  
 CC the detection, quantitation and/or cellular or tissue localisation of a  
 CC polymorphic protein (e.g., for use in measuring levels of the polymorphic  
 CC protein within appropriate physiological samples)  
 XX

SQ Sequence 50 BP; 8 A; 14 C; 18 G; 10 T; 0 U; 0 Other;  
 Query Match 4.2%; Score 20.8; DB 5; Length 50;  
 Best Local Similarity 70.0%; Pred. No. 4.9e+04;  
 Matches 28; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 454 GTGGCCCTTAACATCCCGTCTTGATGTGCTACTGGGGCGGTA 493

OS Synthetic.  
OS Homo sapiens.  
XX WO9833893-A2.  
XX 06-AUG-1998.  
XX 14-JAN-1998; 98WO-US000730.  
XX 31-JAN-1997; 97US-0036476P.  
XX 04-DEC-1997; 97US-00985162.  
XX (RIBO-) RIBOZYME PHARM INC.  
XX (OYAS-) UNIV ASTON.  
XX Akhtar S, Fell P, Mcswiggen JA;  
XX WPI; 1998-437449/37.  
XX Enzymatic nucleic acids - which cleave RNA derived from an epidermal  
PT growth factor receptor, useful for inhibiting cell proliferation and for  
PT treating cancers.  
XX Claim 7; Page 88; 109pp; English.  
XX The present invention describes enzymatic nucleic acid molecules (NMs)  
CC which specifically cleave RNA derived from an epidermal growth factor  
CC receptor (EGF-R) gene. AAV97221 to AAV98043 and AAV98979 to AAV99090  
CC represent specifically claimed target sequence from human EGF-R. AAV98044  
CC to AAV98866 and AAV98867 to V9878 represent hammetthead ribozymes and  
CC hairpin ribozymes respectively for human EGF-R. The NMs are useful for  
CC cleaving EGF-R RNA in the treatment of a condition associated with EGF-R  
CC expression levels e.g. to inhibit cell proliferation in the prevention or  
CC treatment of cancers. The NMs can also be used as diagnostic tools to  
CC examine genetic drift and mutations within diseased cells or to detect  
CC the presence of EGF-R RNA in a cell  
XX  
XX Sequence 50 BP; 16 A; 12 C; 11 G; 0 T; 11 U; 0 Other;

Query Match 4.1%; Score 20.2; DB 2; Length 50;  
Best Local Similarity 60.6%; Pred. No. 7.1e-04;  
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
QY 268 GCACACAGAGTTTCACATGGAAACATGTTG 300  
DB 1 GCACCAAGAGUUUACACAGAAACACACGUG 33

RESULT 14  
ABZ00921/c  
XX ABZ00921 standard; DNA; 50 BP.  
XX ABZ00921;  
XX 09-JAN-2003 (first entry)  
XX Human leukocyte gene expression profiling probe SEQ ID NO 912.  
XX T7; leukocyte; gene expression profiling; allograft rejection;  
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;  
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;  
XX ss.  
XX Homo sapiens.  
XX WO200257414-A2.  
XX 25-JUL-2002.  
XX 22-OCT-2001; 2001WO-US047856.  
XX 20-OCT-2000; 2000US-0241994P.  
XX 08-JUN-2001; 2001US-0296764P.

XX (BIOC-) BIOCARDIA INC.  
XX Wohlgenuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;  
PI Ly N, Woodward R, Quertermous T, Johnson F;  
XX WPI; 2002-636525/68.  
XX New system for leukocyte expression profiling, diagnosing a disease, or  
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis  
PT or congestive heart failure, comprises diagnostic oligonucleotides.  
XX Claim 1; Page 353; Opp; English.  
XX The invention relates to a system for detecting gene expression, which  
CC comprises one or two isolated DNA molecules that detect expression of a  
CC gene, where the gene corresponds to any of 8143 oligonucleotides  
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful  
CC for leukocyte expression profiling. It is particularly useful for  
CC diagnosing a disease, monitoring (rate of) progression of a disease,  
CC predicting therapeutic outcome, determining prognosis for a patient,  
CC predicting disease complications in an individual or monitoring response  
CC to treatment in an individual. The diseases include cardiac allograft  
CC rejection, kidney allograft rejection, liver allograft rejection,  
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,  
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection  
XX  
XX Sequence 50 BP; 11 A; 16 C; 7 G; 16 T; 0 U; 0 Other;  
Query Match 4.0%; Score 20; DB 6; Length 50;  
Best Local Similarity 72.2%; Pred. No. 8.1e+04;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 198 GGGTGCAAGTCTGTGACGGCGTGTGACACAA 233  
DB 41 GGGGCAAGTGTGTGATGACCTGCCAACAAATA 6

RESULT 15  
ABL99142/c  
ID ABL99142 standard; DNA; 40 BP.  
XX ABL99142;  
XX 28-JUN-2002 (first entry)  
XX Green/red click beetle luciferase preparing oligo SEQ ID NO:110.  
XX Luciferase; synthetic nucleic acid; transcriptional characteristic;  
KW transcription; codon usage; PCR; primer; ss.  
XX Coleoptera.  
XX Synthetic.  
XX WO200216944-A2.  
XX 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US026566.  
XX 24-AUG-2000; 2000US-00645706.  
XX (PROM-) PROMEGA CORP.  
XX Wood KV, Wood MG, Zhuang Y, Paguio A;  
XX WPI; 2002-304140/34.  
XX Preparing a synthetic nucleic acid molecule with reduced inappropriate  
PT transcriptional characteristics when expressed in a cell, for e.g making  
PT fusion proteins, by altering a wild type or another synthetic nucleic  
PT acid sequence.  
XX

PS Example 1; Fig 6; 294pp; English.

XX

CC The present invention relates to the preparation of synthetic nucleic

CC acid molecules which have altered transcriptional regulatory sequences

CC compared to the wild-type. These sequences are then transcribed with less

CC frequency compared to the wild-type. In particular, the invention relates

CC to altered luciferase sequences. This can be used to detect weak promoter

CC activity, to express fusion proteins, to detect and/or measure levels of

CC gene expression, subcellular localisation or targeting, in life science

CC research, agrigenetics, gene therapy, developmental science and

CC pharmaceutical development. The present sequence is an oligonucleotide

CC described in the exemplification of the invention

XX

SQ Sequence 40 BP; 13 A; 12 C; 7 G; 8 T; 0 U; 0 Other;

Query Match 4.0%; Score 19.8; DB 6; Length 40;

Best Local Similarity 69.2%; Pred. No. 8.5e-04;

Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 372 TCTGTGTTACTGCTGTGGCTGCGCAGCTTCGGTGAGCT 410

Db 39 TTGGTTGACAAAGTATGATCTGAGCAGCTTGGTGAGCT 1

Search completed: May 26, 2004, 05:53:55

Job time : 429.572 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:16:49 ; Search time 96.1979 Seconds  
(without alignments)  
2872.886 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_1516\_2013

Perfect score: 498  
Sequence: 1 tgcctctgcaaggagcca.....gctactgggcgctaactgc 498

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 839752

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21.2	4.3	50	1	US-08-672-158A-10
2	21	4.2	30	1	Sequence 10, Appl
3	21	4.2	30	1	Sequence 5, Appl
4	21	4.2	30	1	Sequence 5, Appl
5	21	4.2	30	1	Sequence 3, Appl
6	21	4.2	30	1	Sequence 5, Appl
7	21	4.2	30	1	Sequence 3, Appl
8	21	4.2	30	2	Sequence 3, Appl
9	21	4.2	30	2	Sequence 5, Appl
10	21	4.2	30	4	Sequence 3, Appl
11	20.2	4.1	50	3	PCT-US93-11696-5
12	20.2	4.1	50	3	Sequence 5, Appl
13	19.8	4.0	42	2	US-08-985-162-1712
14	19.8	4.0	42	2	Sequence 1712, Ap
15	19.8	4.0	42	2	Sequence 26, Appl
16	19.8	4.0	50	3	Sequence 26, Appl
17	19.6	3.9	48	4	Sequence 1737, Ap
18	19	3.8	36	4	Sequence 20, Appl
19	19	3.8	36	4	Sequence 26, Appl
20	19	3.8	36	5	Sequence 26, Appl
21	19	3.8	42	3	Sequence 4, Appl
22	18.8	3.8	50	1	Sequence 5, Appl
23	18.4	3.7	36	4	Sequence 75, Appl
24	18.4	3.7	36	4	Sequence 25, Appl
25	18.4	3.7	38	2	Sequence 26, Appl
26	18.4	3.7	38	2	Sequence 2015, Ap
27	18.4	3.7	38	2	Sequence 2022, Ap
28	18.4	3.7	38	2	Sequence 2179, Ap

C 28	18.4	3.7	38	3	US-09-071-845-2015	Sequence 2015, Ap
C 29	18.4	3.7	38	3	US-09-071-845-2022	Sequence 2022, Ap
C 30	18.4	3.7	38	3	US-09-071-845-2179	Sequence 2179, Ap
C 31	18.4	3.7	45	4	US-09-765-873A-29	Sequence 29, Appl
C 32	18.4	3.7	45	4	US-09-765-873A-30	Sequence 30, Appl
C 33	18.4	3.7	47	4	US-09-422-878-1087	Sequence 1087, Ap
C 34	18.4	3.7	48	4	US-09-502-558-21	Sequence 21, Appl
C 35	18.4	3.7	48	4	US-09-502-558-22	Sequence 22, Appl
C 36	18.2	3.7	46	3	US-09-372-934-14	Sequence 14, Appl
C 37	18.2	3.7	46	3	US-09-372-934-25	Sequence 25, Appl
C 38	18.2	3.7	46	4	US-09-486-241-18	Sequence 18, Appl
C 39	18.2	3.7	46	4	US-09-766-916-14	Sequence 14, Appl
C 40	18.2	3.7	46	4	US-09-766-916-25	Sequence 25, Appl
C 41	18.2	3.7	46	4	US-09-766-898-14	Sequence 14, Appl
C 42	18.2	3.7	46	4	US-09-766-898-25	Sequence 25, Appl
C 43	18.2	3.7	46	4	US-09-637-040C-14	Sequence 14, Appl
C 44	18.2	3.7	46	4	US-09-637-040C-25	Sequence 25, Appl
C 45	18.2	3.7	50	3	US-08-998-099-283	Sequence 283, App

ALIGNMENTS

RESULT 1  
US-08-672-158A-10/c  
; Sequence 10, Application US/08672158A  
; Patent No. 5770371  
; GENERAL INFORMATION:  
; APPLICANT: Sheryl Thompson  
; TITLE OF INVENTION: Modification of Cryptic Splice Sites In  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 57703710 No. 5770371disk of No. 5770371th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,158A  
; FILING DATE: 27-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Agis Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 4855.000-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 50 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-672-158A-10

Query Match 4.3%; Score 21.2; DB 1; Length 50;  
Best Local Similarity 69.8%; Pred. No. 4.9e+03;  
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qv 396 CAGCTTCGTCGACCTGACCTATCAGTATCGGAGAACATTC 437  
Db 50 CGGAGCGTTCAGCTCGCCGACCATTTATCAGGAGAACATTC 9

RESULT 2

```
US-07-985-692-5
; Sequence 5, Application US/07985692
; Patent No. 5436153
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A
; APPLICANT: Foster, Donald C
; APPLICANT: No. 5436153, Kjel
; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/985,692
; FILING DATE: 19921202
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 92-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ZC4792
;
US-07-985-692-5
Query Match 4.2%; Score 21; DB 1; Length 30;
Best Local Similarity 82.8%; Pred. No. 4.8e+03;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 377 GTTACTGCTGGCTGCGAGCTTCCGT 405
||| ||||| ||||| ||||| |||||
Db 1 GTTGTGCTGTTGCTCGCGAGCTTCCGT 29

RESULT 3
US-08-155-331-5
; Sequence 5, Application US/08155331
; Patent No. 5441931
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy
; APPLICANT: Foster, Donald C
; APPLICANT: No. 5441931, Kjel
; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,710
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 93-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid

US-07-985-692-5
Query Match 4.2%; Score 21; DB 1; Length 30;
Best Local Similarity 82.8%; Pred. No. 4.8e+03;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 377 GTTACTGCTGGCTGCGAGCTTCCGT 405
||| ||||| ||||| ||||| |||||
Db 1 GTTGTGCTGTTGCTCGCGAGCTTCCGT 29

RESULT 4
US-08-147-710-3
; Sequence 3, Application US/08147710
; Patent No. 5455338
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kiesel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,710
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 93-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ZC4792
US-08-147-710-3

Query Match          4.2%; Score 21; DB 1; Length 30;
Best Local Similarity 82.8%; Pred. No. 4.8e+03;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 377 GTTACTGCTGTGGCTGCGCAGCTTCGGT 405
DB 1 GTTGTGCTGTGCTCCGCGAGCTTCGGT 29

RESULT 5
US-08-424-022-5
; Sequence 5, Application US/08424022
; Patent No. 5677146
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy
; APPLICANT: Foster, Donald C
; APPLICANT: No. 5677146ris, Kjeld
; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR
; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,022
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,692
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-684
; REFERENCE/DOCKET NUMBER: 92-21C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ZC4792
US-08-424-022-5

Query Match          4.2%; Score 21; DB 1; Length 30;
Best Local Similarity 82.8%; Pred. No. 4.8e+03;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 377 GTTACTGCTGTGGCTGCGCAGCTTCGGT 405
DB 1 GTTGTGCTGTGCTCCGCGAGCTTCGGT 29

RESULT 6
US-08-458-090-3
```

```
; Sequence 3, Application US/08458090
; Patent No. 5728674
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kiesel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,090
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-14D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ZC4792
US-08-458-090-3

Query Match          4.2%; Score 21; DB 1; Length 30;
Best Local Similarity 82.8%; Pred. No. 4.8e+03;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 377 GTTACTGCTGTGGCTGCGCAGCTTCGGT 405
DB 1 GTTGTGCTGTGCTCCGCGAGCTTCGGT 29

RESULT 7
US-08-457-887-3
; Sequence 3, Application US/08457887
; Patent No. 5914315
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Kiesel, Walter
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS AND
; TITLE OF INVENTION: METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/457,887  
;; FILING DATE:  
;; CLASSIFICATION: 536  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Parker, Gary E  
;; REGISTRATION NUMBER: 31-648  
;; REFERENCE/DOCKET NUMBER: 93-14D1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 206-442-6673  
;; TELEFAX: 206-442-6678  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 30 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; CLONE: ZC4792  
US-08-457-887-3  
  
Query Match 4.2%; Score 21; DB 2; Length 30;  
Best Local Similarity 82.8%; Pred. No. 4.8e+03;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 377 GTTACTGCTGGCGCTCGCAGCTTCGCT 405  
Db 1 GTTGTGCTGTTGCTCGCAGCCTTCGCT 29  
  
RESULT 8  
US-08-424-017B-5  
; Sequence 5, Application US/08424017B  
; Patent No. 5935854  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Donald C  
; APPLICANT: Sprecher, Cindy  
; APPLICANT: Kiesel, Walter  
; APPLICANT: Foster, Donald C  
; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR  
; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 4225 Roosevelt Way, N.E.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/155,331  
; FILING DATE:  
; APPLICATION NUMBER: US 07/985,692  
; FILING DATE: 02-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, Gary E  
; REGISTRATION NUMBER: 31-684  
; REFERENCE/DOCKET NUMBER: 92-21C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-547-8080 ext 322  
; TELEFAX: 206-548-2329  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 30 base pairs

;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; CLONE: ZC4792  
US-08-424-017B-5  
  
Query Match 4.2%; Score 21; DB 2; Length 30;  
Best Local Similarity 82.8%; Pred. No. 4.8e+03;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 377 GTTACTGCTGGCGCTCGCAGCTTCGCT 405  
Db 1 GTTGTGCTGTTGCTCGCAGCCTTCGCT 29  
  
RESULT 9  
US-09-904-621-3  
; Sequence 3, Application US/09904621  
; Patent No. 6656746  
; GENERAL INFORMATION:  
; APPLICANT: Sprecher, Cindy A.  
; APPLICANT: Kiesel, Walter  
; APPLICANT: Foster, Donald C  
; TITLE OF INVENTION: NOVEL HUMAN KUNITZ-TYPE INHIBITORS  
; TITLE OF INVENTION: AND  
; TITLE OF INVENTION: METHODS RELATING THERETO  
; FILE REFERENCE: 93-14D3  
; CURRENT APPLICATION NUMBER: US/09/904,621  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/265,627  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 5,455,338  
; PRIOR FILING DATE: EARLIER FILING DATE: 1993-11-05  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 30  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Oligonucleotide ZC4792  
US-09-904-621-3  
  
Query Match 4.2%; Score 21; DB 4; Length 30;  
Best Local Similarity 82.8%; Pred. No. 4.8e+03;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
QY 377 GTTACTGCTGGCGCTCGCAGCTTCGCT 405  
Db 1 GTTGTGCTGTTGCTCGCAGCCTTCGCT 29  
  
RESULT 10  
PCT-US93-11696-5  
; Sequence 5, Application PC/TUS9311696  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Donald C  
; APPLICANT: Sprecher, Cindy  
; APPLICANT: Norris, Kjeld  
; TITLE OF INVENTION: NOVEL HUMAN AMYLOID PROTEIN PRECURSOR  
; TITLE OF INVENTION: HOMOLOG AND KUNITZ-TYPE INHIBITOR  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 4225 Roosevelt Way, N.E.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible



```

/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 1712:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 50 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
US-08-985-162-1712
Query Match 4.1%; Score 20.2; DB 3; Length 50;
Best Local Similarity 60.6%; Pred. No. 9.3e+03;
Matches 20; Conservative 5; Mismatches 8; Indels

QY 268 GCACCCGAGCGTTTGACATCGAAAAACATGTTG 300
DB 1 GCACCCGAGAGUUTUACCGAGAGAACACACGUG 33

RESULT 12
US-09-401-063-1712
/ Sequence 1712, Application US/09401063
/ Patent No. 6623962
/ GENERAL INFORMATION:
/ APPLICANT: Akhtar, Saghir
/ APPLICANT: Fell, Patricia
/ APPLICANT: McSwigen, James
/ TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT
/ TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
/ TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
/ TITLE OF INVENTION: FACTOR RECEPTORS
/ NUMBER OF SEQUENCES: 1877
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ STREET: Suite 4700
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071-2066
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ MEDIUM TYPE: storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: FastSeq for Windows 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/401,063
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/985,162
/ FILING DATE: 04 December 1997
/ APPLICATION NUMBER: 60/036,476
/ FILING DATE: 31 January 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Warburg, Richard J.
/ REGISTRATION NUMBER: 32,327
/ REFERENCE/DOCKET NUMBER: 230/107
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 1712:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 50 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
US-09-401-063-1712

```

```

Query Match      4.1%; Score 20.2; DB 4; Length 50;
Best Local Similarity 60.6%; Pred. No. 9.3e+03;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY      268 GCACACAGAGGTTTCACATGGGAAAACATGTTG 300
      |||||:::|||||||:::|
Db       1 GCACCCAGAAGUUUACACAGAAACACACGUUG 33

RESULT 13
US-09-133-774-26
; Sequence 26, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu W.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636el Peptides Capable of Modulating Inflammatory Heart
; FILE REFERENCE: Disease
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Human
US-09-133-774-26

```

Query Match	4.0%;	Score 19.8;	DB 2;	Length 42;
Best Local Similarity	77.4%;	Pred. No. 1.1e-04;		
Matches 24;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
QY	115	CTGCAGCCCTTTGTGCCA	CTGTACTGGGGCT	145
Db	12	CAGCACCCCTGTTGGCAA	CTATCTGGGGCT	42

RESULT 14  
US-09-303-862-26  
; Sequence 26, Application US/09303862  
; Patent No. 6034230  
; GENERAL INFORMATION:  
; APPLICANT: Bachmaier, Kurt  
; APPLICANT: Hessel, Andrew J.  
; APPLICANT: Neu M.D., Nikolaus  
; APPLICANT: Penninger, Josef M.  
; TITLE OF INVENTION: No. 6034230el Peptides Capable of Modulating Inflammatory Heart  
; TITLE OF INVENTION: Disease  
; FILE REFERENCE: A-536  
; CURRENT APPLICATION NUMBER: US/09/303,862  
; CURRENT FILING DATE: 1999-05-03  
; EARLIER APPLICATION NUMBER: 09/133,774  
; EARLIER FILING DATE: 1998-08-12  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 42  
; TYPE: DNA  
; ORGANISM: Human  
US-09-303-862-26

```

Query Match          4.0%; Score 19.8; DB 3; Length 42;
Best Local Similarity 77.4%; Pred. No. 1.1e+04;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Qy 115 CTGCAGCCTTCTCTGCCACCTGTACTGGGCT 145

Db 12 CAGCACCTGTTTGCCAACTATGCTGGGCT 42

RESULT 15  
 US-08-985-162-1737  
 ; Sequence 1737, Application US/08985162  
 ; Patent No. 6057156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Akhtar, Saghir  
 ; APPLICANT: Fell, Patricia  
 ; APPLICANT: McSwigen, James  
 ; TITLE OF INVENTION: ENZYMAIC NUCLEIC ACID TREATMENT  
 ; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED  
 ; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH  
 ; TITLE OF INVENTION: FACTOR RECEPTORS  
 ; NUMBER OF SEQUENCES: 1877  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071-2066  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: Storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: FastSeq for Windows 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/985,162  
 ; FILING DATE: 04 December 1997  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/036,476  
 ; FILING DATE: 31 January 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warburg, Richard J.  
 ; REGISTRATION NUMBER: 32,327  
 ; REFERENCE/DOCKET NUMBER: 230/107  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 1737:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 50 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 US-08-985-162-1737

Query Match	4.0%	Score 19.8	DB 3	Length 50;
Best Local Similarity	51.1%	Pred. No. 1.2e+04;		
Matches 24;	Conservative	6;	Mismatches 17;	Indels 0;
				Caps 0;

QY 224 TGAACAACAGCTACGAGTCAGACATCCTGAAGAATTACCTGGCA 270  
:  
DB 4 UGAAGAAGAACACCAGAGAAACACACCGTUGUGGUA CAU AACCUUGUA 50

Search completed: May 26, 2004, 09:20:42  
Job time : 97.1979 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:41:55 ; Search time 2173.84 Seconds  
(without alignments)  
1041.142 Million cell updates/sec

Title: US-10-048-046-1\_COPY1516\_2013  
Perfect score: 498  
Sequence: 1 tgcctctgcaaggaagca.....gctactggggcgtaactgc 498

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 1612472

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	4.2	30	9	US-09-904-621-3
2	20.8	4.2	48	12	US-10-251-085B-26
3	20.8	4.2	48	12	US-10-251-085B-64
4	20.4	4.1	47	16	US-10-415-247-6
5	20	4.0	50	16	US-10-131-827-912
6	19.6	3.9	48	11	US-09-767-764A-12
7	19.6	3.9	48	15	US-10-285-010-20
8	19.6	3.9	50	16	US-10-131-827-7822
9	19.4	3.9	39	15	US-10-173-461-24
10	19	3.8	36	15	US-10-008-960-26
11	18.8	3.8	48	12	US-10-251-085B-27
12	18.8	3.8	48	12	US-10-251-085B-38
13	18.8	3.8	48	12	US-10-251-085B-65
14	18.8	3.8	48	12	US-10-251-085B-76

APPLICANT: LY, NGOC  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES

```
; Sequence 12, Application US/09767764A
; Publication No. US20030207265A1
; GENERAL INFORMATION:
; APPLICANT: Church, George M.
; TITLE OF INVENTION: METHOD OF MAKING PROTEIN ARRAYS
; FILE REFERENCE: 10498-00009
; CURRENT APPLICATION NUMBER: US/09767,764A
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/522,732
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Amplification primer
US-09-767-764A-12

Query Match          3.9%; Score 19.6; DB 11; Length 48;
Best Local Similarity 73.5%; Pred. No. 3.5e+04;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 36 CTGCTCCAGCCCATGCCCGACGGAGCGGAG 69
DB 43 CTCCTCAAGCTTCGCCCGCAGAGGAAAGCGGAG 10

RESULT 7
US-10-285-010-20/c
; Sequence 20, Application US/10285010
; Publication No. US20030124594A1
; GENERAL INFORMATION:
; APPLICANT: Church, George
; APPLICANT: Mitra, Robi
; TITLE OF INVENTION: Replica Amplification Of Nucleic Acid Arrays
; FILE REFERENCE: 10498-00040
; CURRENT APPLICATION NUMBER: US/10/285,010
; CURRENT FILING DATE: 2002-10-31
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amplification primer
US-10-285-010-20

Query Match          3.9%; Score 19.6; DB 15; Length 48;
Best Local Similarity 73.5%; Pred. No. 3.5e+04;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 36 CTGCTCCAGCCCATGCCCGACGGAGCGGAG 69
DB 43 CTCCTCAAGCTTCGCCCGCAGAGGAAAGCGGAG 10

RESULT 8
US-10-131-827-7822
; Sequence 7822, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fty, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
```

```
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7822
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-7822

Query Match          3.9%; Score 19.6; DB 16; Length 50;
Best Local Similarity 66.7%; Pred. No. 3.6e+04;
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 197 TGGGTGACAAAGTCTCTGGACGGCGTGTGAACAACAACAGCT 238
DB 1 TGGCGAGCCAGTCTCTGGATGGGATTCGTGATCAACAGAGTT 42

RESULT 9
US-10-173-461-24
; Sequence 24, Application US/10173461
; Publication No. US20030138795A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN GROWTH FACTOR WITH HOMOLOGY
; TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR, BGS-8, EXPRESSED HIGHLY IN IMMUNE TISSUE
; FILE REFERENCE: D0166 NP
; CURRENT APPLICATION NUMBER: US/10/173,461
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/298,340
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 24
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-461-24

Query Match          3.9%; Score 19.4; DB 15; Length 39;
Best Local Similarity 70.3%; Pred. No. 3.9e+04;
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 158 GCTGCTACGGCTGCTGGCCCGCTTTTGTGAGCTCAA 194
DB 1 GCAGCAGCGCGCGCATGGCCCTCTGGAGTGAGCACCA 37

RESULT 10
US-10-008-960-26
; Sequence 26, Application US/10008960
; Publication No. US2003013971A1
; GENERAL INFORMATION:
; APPLICANT: BAYLOR COLLEGE OF MEDICINE
; APPLICANT: SMITH, JAMES R.
; APPLICANT: DRUTZ, DAVID J.
; APPLICANT: WILSON, DEBORAH R.
; APPLICANT: ZUMSTEIN, LOUIS A.
; TITLE OF INVENTION: SENESENT CELL DERIVED INHIBITORS OF
; TITLE OF INVENTION: DNA SYNTHESIS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROGERS & WELLS
; STREET: 200 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10166
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21
; CURRENT APPLICATION NUMBER: US/10/251,085B
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 65
; LENGTH: 48
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
US-10-251-085B-65

Query Match      3.8%; Score 18.8; DB 12; Length 48;
Best Local Similarity 65.0%; Pred. No. 6.4e+04;
Matches 26; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 50 TGGCCGTTGGAGAGGAGTGCTCGAGSAGTCCAGCTGGT 89
Db 5 TGGCCGTTGGAGAGGAGTGCTCGAGSAGTCCAGCTGGT 44

RESULT 14
US-10-251-085B-76
; Sequence 76, Application US/10251085B
; Publication No. US20040072164A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Maruyama, Toshiaki
; TITLE OF INVENTION: ENGINEERED TEMPLATES AND THEIR USE IN SINGLE PRIMER AMPLIFICATION
; FILE REFERENCE: 1087-21
; CURRENT APPLICATION NUMBER: US/10/251,085B
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 60/323,455
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 76
; LENGTH: 48
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
US-10-251-085B-76

Query Match      3.8%; Score 18.8; DB 12; Length 48;
Best Local Similarity 65.0%; Pred. No. 6.4e+04;
Matches 26; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 50 TGGCCGTTGGAGAGGAGTGCTCGAGSAGTCCAGCTGGT 89
Db 5 TGGCCGTTGGAGAGGAGTGCTCGAGSAGTCCAGCTGGT 44

RESULT 15
US-09-930-423-3110/c
; Sequence 3110, Application US/09930423
; Publication No. US20030092003A1
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: MEHB00,918-A 400/027
; CURRENT APPLICATION NUMBER: US/09/930,423
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 4553
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3110
; LENGTH: 38
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-930-423-3110

Query Match      3.7%; Score 18.6; DB 10; Length 38;
Best Local Similarity 72.7%; Pred. No. 6.9e+04;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 64 GCGGAGCGCGAGCAGGACCCGCGTGTCGCCCT 96
Db 38 GCGCAGCGAGACCTTGACTCGCCTTTGGGCCT 6

Search completed: May 26, 2004, 10:22:40
Job time : 2173.84 secs
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:12:53 ; Search time 3457.26 Seconds  
(without alignments)

4301.492 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_1516\_2013

Perfect score: 498

Sequence: 1 tgcctctgcaaggaagcca.....gctactgggcgtaactgc 498

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rtd:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.8	4.2	50	9	AU105995
2	20.8	4.2	50	9	AU105997
3	20.6	4.1	50	9	AU105996
4	20.2	4.1	47	28	CC178014

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	19.6	3.9	43	9	AA974059
C	6	19.6	3.9	50	9	AU103246
C	7	19.4	3.9	49	28	AZ606700
C	8	19.2	3.9	49	9	AA990168
C	9	19	3.8	49	9	AU1271755
C	10	18.8	3.8	50	9	AU105998
C	11	18.6	3.7	40	10	BF343276
C	12	18.6	3.7	49	9	AA921726
C	13	18.6	3.7	50	9	AU103244
C	14	18.6	3.7	50	9	AU104470
C	15	18.6	3.7	50	9	AU105005
C	16	18.6	3.7	50	12	BI665523
C	17	18.4	3.7	46	9	AI949114
C	18	18.4	3.7	50	9	AU105237
C	19	18.4	3.7	50	9	AU107915
C	20	18.2	3.7	48	29	CG779308
C	21	18.2	3.7	49	9	AA933611
C	22	18	3.6	37	9	AI048978
C	23	18	3.6	43	12	BI103457
C	24	18	3.6	46	28	AZ512864
C	25	18	3.6	50	9	AU103249
C	26	18	3.6	50	9	AU105753
C	27	18	3.6	50	9	AU105754
C	28	18	3.6	50	9	AU105755
C	29	18	3.6	50	9	AU107902
C	30	18	3.6	50	9	AU107933
C	31	17.8	3.6	38	28	AZ868035
C	32	17.8	3.6	46	9	AA879784
C	33	17.8	3.6	46	28	CC183613
C	34	17.8	3.6	49	28	BH627713
C	35	17.8	3.6	50	9	AU102353
C	36	17.8	3.6	50	9	AU102354
C	37	17.8	3.6	50	9	AU105314
C	38	17.8	3.6	50	9	AU106718
C	39	17.6	3.5	46	9	AI688909
C	40	17.6	3.5	46	9	AA410197
C	41	17.6	3.5	49	9	AI001027
C	42	17.6	3.5	49	12	BM662480
C	43	17.6	3.5	50	9	AU103245
C	44	17.6	3.5	50	9	AU104468
C	45	17.6	3.5	50	9	AU105698

## ALIGNMENTS

RESULT 1	AU105995	50 bp	mrna	linear	EST 30-AUG-2001
LOCUS	AU105995	Sugano Homo sapiens cDNA library	Homo sapiens	cdna	clone
DEFINITION	ADSE00435, mRNA sequence.				
ACCESSION	AU105995				
VERSION	AU105995.1	GI:13555516			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.				
TITLE	Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites				
JOURNAL	EMBO Rep. 2 (5), 388-393 (2001)				
MEDLINE	21270072				
PUBMED	11375929				
COMMENT	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and				



Accession	LOCUS	Size	Library	EST
U05996	LOCUS	50 bp	mRNA	EST 30-AUG-2001
U05996	DEFINITION		Sugano Homo sapiens cDNA library	
U05996	DEFINITION		Homo sapiens cDNA library	
U05996	DEFINITION		Homo sapiens cDNA clone	

```
/organism="Mus musculus"
```

/mol\_type="mRNA"  
/strain="129 OLA"  
/db\_xref="taxon:10090"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/clone\_lib="BayGenomics Gene Trap Library pGTLxlf"  
/note="Vector: pGTLxlf"

ORIGIN

Query Match 4.1%; Score 20.2; DB 28; Length 47;  
Best Local Similarity 68.3%; Pred. No. 8.9e+05;  
Matches 28; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 113 GCTCGAGCCTTCTGCCACCTGTACTGGGGTGCACCCGG 153  
|||  
Db 7 GCTTCTGCTTTTCGCCCTCCGAGATGCGCTGGACCCCG 47

RESULT 5  
AA974059/c  
LOCUS  
DEFINITION  
OQ05h07.s1 NCI CGAP Lu5 Homo sapiens cDNA clone IMAGE:1585501.3;  
similar to SW/ASHI HUMAN P50553 ACHAETE-SCUTE HOMOLOG 1.; contains  
element TAK1 repetitive element 1.; mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabsx@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality  
Insert length: 802 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers  
1..43  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1585501"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI-CGAP Lu5"  
/notes="Organ: lung; Vector: p773D-Pac (Pharmacia) with a  
modified polylinker; 1st strand cDNA was prepared from  
neuroendocrine lung carcinoid, and was then primed with a  
Not I - oligo(dT) primer. Double-stranded cDNA was ligated  
to Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
p773 vector. Library is normalized. Library was  
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 3.9%; Score 19.6; DB 9; Length 43;  
Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 146 GCACCCGACCGCTGCTACGGTGTGCTGCCCGCTTTGTG 187  
|||  
Db 42 GCCCCTGAGCGGTGGTGGCGCGCGTCTCAGCTCGCG 1

RESULT 6  
AU103246/c  
LOCUS  
DEFINITION  
AU103246 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
KAT01774, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yezukui@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and  
Sugano, S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),  
149-156 (1997).  
Location/Qualifiers  
1..50  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="KAT01774"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 3.9%; Score 19.6; DB 9; Length 50;  
Best Local Similarity 66.7%; Pred. No. 1.3e+06;  
Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 40 TTCAGCCCATGCCGACCGAGCGGAGCGGAGCGGAGCAGGAC 81  
|||  
Db 45 TACCAGACCGCGCCCTTGTGTGAGACGCGCGCGGAGCCGGAC 4

RESULT 7  
AZ606700/c  
LOCUS  
DEFINITION  
AZ606700 Mouse 10kb plasmid UGCLM library Mus musculus genomic  
clone UUGC1M0428E15 R, genomic survey sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
TITLE  
Niederhausen, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

[illegible]

cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality  
Insert Length: 1806 Std Error: 0.00  
Seq primer: -40up from Gibco  
High quality sequence stop: 1.  
Location/Qualifiers

#### FEATURES

source

1. .49

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1865484"

/lab\_host="DH10B"

/clone\_lib="NCI\_CGAP Kid3"

/note="Organ: kidney; Vector: p773D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer,  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified p773 vector. mRNA  
source: 2 pooled kidneys. Library went through one round  
of normalization. Library constructed by Bento Soares and  
M. Fatima Bonaldo. "

#### ORIGIN

Query Match 3.8%; Score 19; DB 9; Length 49;

Best Local Similarity 71.4%; Pred. No. 1.8e+06;

Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 62 GAGCGAGCGCGAGGAGCGCGGTGCGCCCT 96

Db 39 GGCGGCACTCCAGCAGCGCGCGCTGCTGCT 5

#### RESULT 10

AU105998

LOCUS AU105998 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

W0521, mRNA sequence.

DEFINITION AU105998

VERSION AU105998.1 GI:13555519

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,

Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K.,

Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.

Diverse transcriptional initiation revealed by fine, large-scale

mapping of mRNA start sites

JOURNAL ENBO Rep. 2 (5), 388-393 (2001)

MEDLINE 21270072

PUBMED 11375929

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and

Sugano, S. Construction and characterization of a full

length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),

149-156 (1997).

#### FEATURES

source

1. .50

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="W0521"

/clone\_lib="Sugano Homo sapiens cDNA library"

#### ORIGIN

Query Match 3.8%; Score 18.8; DB 9; Length 50;

Best Local Similarity 63.0%; Pred. No. 2e+06;

Matches 29; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 38 GCTTCCAGCCCATGCCCGACCGGAGCGGAGCGCGAGCGAGCGGCC 83

Db 1 GTTGGCTCCGAGGCTGGCGCGGAGCGTGGAGCCCGCGCGCTGCC 46

#### RESULT 11

BF343276/c

LOCUS BF343276

DEFINITION 602015935F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4151432

5', mRNA sequence.

ACCESSION BF343276

VERSION BF343276.1 GI:11290436

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 40)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM9416 row: f column: 09

High quality sequence stop: 38.

#### FEATURES

source

1. .40

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4151432"

/tissue\_type="glioblastoma with EGFR amplification"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP Brn64"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.57 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

#### ORIGIN

Query Match 3.7%; Score 18.6; DB 10; Length 40;

Best Local Similarity 72.7%; Pred. No. 2e+06; 9; Indels 0; Gaps 0;

Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 140 GGGGCTGCACCCGAGCGCGCTGTCTACGCGTCC 172

Db 33 GGAGCTGCAGACGACCGCTGGGAGCCGCGCCGAGCC 1

#### RESULT 12

AA921726/c

LOCUS AA921726

DEFINITION AA921726 49 bp mRNA linear EST 13-MAY-1998

oh08d02.s1 NCI\_CGAP C08 Homo sapiens cDNA clone IMAGE:1457187 3,

similar to TR:O15551 GI5551 RAT VENTRAL PROSTATE.1 HOMOLOG. ;, mRNA

sequence.

ACCESSION AA921726

VERSION AA921726.1 GI:3068914



Search completed: May 26, 2004, 09:17:50  
Job time : 3459.26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:10:28 ; Search time 1630.25 Seconds  
(without alignments)  
5822.506 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_181\_399  
Perfect score: 219  
Sequence: 1 gctctctgaggaagcgga.....tacagactgggagtgatc 219

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pri:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hg\_hum:\*  
31: em\_hg\_inv:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pln:\*  
35: em\_hg\_rod:\*  
36: em\_hg\_nam:\*  
37: em\_hg\_vrt:\*  
38: em\_sy:\*  
39: em\_hg\_hum:\*  
40: em\_hg\_mus:\*  
41: em\_hg\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	19.8	9.0	50	6	AX165857	AX165857 Sequence
2	19.4	8.9	40	6	AB3617	AB3617 Sequence 46
3	19	8.7	47	6	AX194693	AX194693 Sequence
4	18.8	8.6	38	6	AR013934	AR013934 Sequence
5	18.4	8.4	38	6	AR020204	AR020204 Sequence
6	18.4	8.4	38	6	AR032156	AR032156 Sequence
7	18.4	8.4	38	6	AR043821	AR043821 Sequence
8	18.4	8.4	38	6	AR072133	AR072133 Sequence
9	18.4	8.4	38	6	AR095987	AR095987 Sequence
10	18.4	8.4	38	6	AR136647	AR136647 Sequence
11	18.4	8.4	38	6	AR157182	AR157182 Sequence
12	18.4	8.4	38	6	AR164131	AR164131 Sequence
13	18.4	8.4	38	6	I32817	I32817 Sequence 13
14	18.4	8.4	38	6	I59631	I59631 Sequence 13
15	18.4	8.4	38	6	AX127254	AX127254 Sequence
16	18.4	8.4	38	6	AX800785	AX800785 Sequence
17	18.4	8.4	46	6	I18512	I18512 Sequence 12
18	18.4	8.4	47	6	AR289961	AR289961 Sequence
19	18.2	8.3	43	6	BD174678	BD174678 Ribozyme
20	18	8.2	29	6	BD140120	BD140120 Secreted
21	18	8.2	42	6	AX328821	AX328821 Sequence
22	18	8.2	42	6	BD132386	BD132386 DNA diag
23	18	8.2	50	6	E44269	E44269 Oligo-DNA s
24	18	8.2	50	6	AX697058	AX697058 Sequence
25	17.8	8.1	49	6	AR032442	AR032442 Sequence
26	17.8	8.1	49	6	I29182	I29182 Sequence 54
27	17.8	8.1	49	6	I90856	I90856 Sequence 54
28	17.8	8.1	49	6	AR209106	AR209106 Sequence
29	17.6	8.0	40	6	AX060727	AX060727 Sequence
30	17.6	8.0	40	6	AX060906	AX060906 Sequence
31	17.6	8.0	41	6	AX518097	AX518097 Sequence
32	17.4	7.9	41	6	AX513875	AX513875 Sequence
33	17.4	7.9	41	6	AX519151	AX519151 Sequence
34	17.4	7.9	47	6	AR284501	AR284501 Sequence
35	17.4	7.9	50	6	AR032934	AR032934 Sequence
36	17.4	7.9	50	6	I29674	I29674 Sequence 54
37	17.4	7.9	50	6	I91348	I91348 Sequence 54
38	17.4	7.9	50	6	AR209598	AR209598 Sequence
39	17.2	7.9	38	6	AR210306	AR210306 Sequence
40	17.2	7.9	38	6	AX402734	AX402734 Sequence
41	17.2	7.9	41	6	AR061380	AR061380 Sequence
42	17.2	7.9	41	6	AR108279	AR108279 Sequence
43	17.2	7.9	41	6	I16236	I16236 Sequence 82
44	17.2	7.9	41	6	I66722	I66722 Sequence 62
45	17.2	7.9	41	6	I84816	I84816 Sequence 62

ALIGNMENTS

RESULT 1  
AX165857  
LOCUS AX165857 50 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 1052 from Patent WO0138586.  
ACCESSION AX165857  
VERSION AX165857.1 GI:14546686  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Shimkets R.A. and Leach M.  
TITLE Nucleic acids containing single nucleotide polymorphisms and  
Methods of use thereof

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0138586-A 1052 31-MAY-2001;  
Curagen Corporation (US)  
FEATURES Location/Qualifiers  
source 1..50  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
misc\_feature 25..26  
/note="Nucleotide deleted between bases 25 and 26"  
Accession number CG4004690  
variation 26  
/note="single nucleotide polymorphism"  
ORIGIN  
Query Match 9.0%; Score 19.8; DB 6; Length 50;  
Best Local Similarity 69.2%; Pred. No. 2.3e+05;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 9 GAGGAGCGGAGGTGGACCATCGGCGGAGGAGGAGGTG 47  
Db 1 GAGGAGGAGGAGGTGGAGGAGGAGGAGGAGGAGGATG 39  
RESULT 2  
LOCUS A83617 40 bp DNA linear PAT 21-JAN-2000  
DEFINITION Sequence 46 from Patent WO9849324.  
ACCESSION A83617  
VERSION A83617.1 GI:6732869  
KEYWORDS unidentifed  
SOURCE unidentifed  
ORGANISM unclassified  
REFERENCE 1 (bases 1 to 40)  
AUTHORS Matthijs,G.  
TITLE CARBOHYDRATE-DEFICIENT GLYCOPROTEIN SYNDROME TYPE I  
JOURNAL Patent: WO 9849324-A 46 05-NOV-1998;  
MATTHIJS GERT (BE); GENZYME LTD (GB)  
FEATURES Location/Qualifiers  
source 1..40  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 8.9%; Score 19.4; DB 6; Length 40;  
Best Local Similarity 79.3%; Pred. No. 3.1e+05;  
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 97 ATTGTAGTGGATGAATAATCAGGTCAGGT 125  
Db 4 ATTGTAGTGGTGAATAAGTGTAGACAGGT 32  
RESULT 3  
AX194693/c  
LOCUS AX194693 47 bp DNA linear PAT 28-AUG-2001  
DEFINITION Sequence 163 from Patent WO0151659.  
ACCESSION AX194693  
VERSION AX194693.1 GI:15385340  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Chu,T., Blumenfeld,M. and Cohen,D.  
TITLE Biallelic markers derived from genomic regions carrying genes involved in central nervous system disorders  
JOURNAL Patent: WO 0151659-A 163 19-JUL-2001;  
GENSET (FR)  
FEATURES Location/Qualifiers  
source 1..47  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
variation 24  
/note="99-27562-366 : polymorphic base G or T"  
ORIGIN  
Query Match 8.7%; Score 19; DB 6; Length 47;  
Best Local Similarity 67.6%; Pred. No. 4.2e+05;  
Matches 25; Conservative 1; Mismatches 11; Indels 0; Gaps 0;  
QY 170 TGAAGGTTGTTAAGAAGCAGACATGCCCTTTACAGAC 206  
Db 38 TGTAGATGTTTAAAGCATTCCTGGCCTCTACCCAC 2  
RESULT 4  
AR013934/c  
LOCUS AR013934 38 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 23 from patent US 5773223.  
ACCESSION AR013934  
VERSION AR013934.1 GI:3971388  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 38)  
AUTHORS Shyamala,V. and Olson,P.Tekamp.  
TITLE Endothelin B sub.1, (ETB sub.1) receptor polypeptide and its encoding nucleic acid methods, and uses thereof  
JOURNAL Patent: US 5773223-A 23 30-JUN-1998;  
FEATURES Location/Qualifiers  
source 1..38  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 8.6%; Score 18.8; DB 6; Length 38;  
Best Local Similarity 68.4%; Pred. No. 4.9e+05;  
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 65 GCATTAACCTGCTCTCGAGATCAGTGTAGAAATTGTA 102  
Db 38 GCAACAGCTCGATATCTCGAAGATAAAATAGAAATTGTA 1  
RESULT 5  
AR020204/c  
LOCUS AR020204 38 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 13 from patent US 5789156.  
ACCESSION AR020204  
VERSION AR020204.1 GI:3974819  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 38)  
AUTHORS Bujard,H. and Gossen,M.  
TITLE Tetracycline-regulated transcriptional inhibitors  
JOURNAL Patent: US 5789156-A 13 04-AUG-1998;  
FEATURES Location/Qualifiers  
source 1..38  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 8.4%; Score 18.4; DB 6; Length 38;  
Best Local Similarity 69.4%; Pred. No. 6.7e+05;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 148 AGTGGACAGTGTATTAAACAAGCTGAAGGTTGTTAAG 183  
Db 38 AATTAACTGTGATAAAGTAACTAGCTTATCGATGATAG 3



TITLE Mice transgenic for a tetracycline-inducible transcriptional activator  
JOURNAL Patent: US 5912411-A 13 15-JUN-1999;  
FEATURES Location/Qualifiers  
source 1..38  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 8.4%; Score 18.4; DB 6; Length 38;  
Best Local Similarity 69.4%; Pred. No. 6.7e+05;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGAACAGTGATTAAACAAGCTGAAGGTGTGTTAAG 183  
Db 38 AATTAACTGTGATTAACCTAGCTTATCGATGATAAG 3

RESULT 9  
AR095987/c  
LOCUS 38 bp DNA  
DEFINITION Sequence 13 from patent US 6004941.  
ACCESSION AR095987  
VERSION AR095987.1 GI:10024380  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 38)  
AUTHORS Bujard,H. and Gossen,M.  
TITLE Methods for regulating gene expression  
JOURNAL Patent: US 6004941-A 13 21-DEC-1999;  
FEATURES Location/Qualifiers  
source 1..38  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 8.4%; Score 18.4; DB 6; Length 38;  
Best Local Similarity 69.4%; Pred. No. 6.7e+05;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGAACAGTGATTAAACAAGCTGAAGGTGTGTTAAG 183  
Db 38 AATTAACTGTGATTAACCTAGCTTATCGATGATAAG 3

RESULT 10  
AR136647/c  
LOCUS 38 bp DNA  
DEFINITION Sequence 13 from patent US 6136954.  
ACCESSION AR136647  
VERSION AR136647.1 GI:14477319  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 38)  
AUTHORS Bujard,H. and Gossen,M.  
TITLE Tetracycline-inducible transcriptional activator fusion proteins  
JOURNAL Patent: US 6136954-A 13 24-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..38  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 8.4%; Score 18.4; DB 6; Length 38;  
Best Local Similarity 69.4%; Pred. No. 6.7e+05;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGAACAGTGATTAAACAAGCTGAAGGTGTGTTAAG 183

AUTHORS Bujard,H., Gossen,M., Hillen,W., Helbl,V. and Schnappinger,D.  
 TITLE Tetracycline-regulated transcriptional modulators with altered DNA binding specificities  
 JOURNAL Patent: US 5589362-A 13 31-DEC-1996;  
 FEATURES Location/Qualifiers  
 source 1..38  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 8.4%; Score 18.4; DB 6; Length 38;  
 Best Local Similarity 69.4%; Pred. No. 6.7e+05;  
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAAACAAGCTGAAGGTTGTTAAG 183  
 Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 14  
 I59631/c  
 LOCUS 159631 38 bp DNA linear PAT 07-OCT-1997  
 DEFINITION Sequence 13 from patent US 5654168.  
 ACCESSION I59631  
 VERSION I59631.1 GI:2478263  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 38)  
 AUTHORS Bujard,H. and Gossen,M.  
 TITLE Tetracycline-inducible transcriptional activator and tetracycline-regulated transcription units  
 JOURNAL Patent: US 5654168-A 13 05-AUG-1997;  
 FEATURES Location/Qualifiers  
 source 1..38  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 8.4%; Score 18.4; DB 6; Length 38;  
 Best Local Similarity 69.4%; Pred. No. 6.7e+05;  
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGAACAGTGATTAAACAAGCTGAAGGTTGTTAAG 183  
 Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 15  
 AX127254/c  
 LOCUS AX127254 38 bp DNA linear PAT 15-MAY-2001  
 DEFINITION Sequence 13 from Patent EP1092771.  
 ACCESSION AX127254  
 VERSION AX127254.1 GI:14133334  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 REFERENCE 1  
 AUTHORS Bujard,H. and Gossen,M.  
 TITLE Tetracycline-regulated transcriptional modulators  
 JOURNAL Patent: EP 1092771-A 13 18-APR-2001;  
 BASF AKTIENGESSELLSCHAFT (DE); KNOLL AKTIENGESSELLSCHAFT (DE)  
 FEATURES Location/Qualifiers  
 source 1..38  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN

Query Match 8.4%; Score 18.4; DB 6; Length 38;  
 Best Local Similarity 69.4%; Pred. No. 6.7e+05;

Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 148 AGTGGACAGTGTATTAACTAAGCTGAAGTTGTTAAG 183  
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

Search completed: May 26, 2004, 07:39:16  
Job time : 1632.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:01:38 ; Search time 184.951 Seconds  
(without alignments)  
5030.293 Million cell updates/sec

Title: US-10-048-046-1\_COPY181\_399

Perfect score: 219  
Sequence: 1 gtctctctggaagcggga.....tacagactggggtgcatc 219

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002s.\*  
7: Geneseq2003as.\*  
8: Geneseq2003bs.\*  
9: Geneseq2003cs.\*  
10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20.2	9.2	50	4	AAL29957	Human SNP
C 2	20	9.1	28	5	AAL30355	Human che
C 3	19.8	9.0	50	5	ABL01061	Human SNP
C 4	19.2	8.8	41	6	ABV76434	Human DNA
C 5	19	8.7	47	4	AAH88323	CNS disor
C 6	18.8	8.6	38	2	AAV35018	Human end
C 7	18.8	8.6	47	3	AZ67349	Human map
C 8	18.4	8.4	38	2	AAT11362	Class C t
C 9	18.4	8.4	38	2	AAT45713	Class C t
C 10	18.4	8.4	38	2	AAV60085	Class C t
C 11	18.4	8.4	38	2	AAV60051	Class C t
C 12	18.4	8.4	38	3	AZ56132	Class C t
C 13	18.4	8.4	38	4	AAZ61572	Class C t
C 14	18.4	8.4	38	4	AAH47637	Nucleotid
C 15	18.4	8.4	38	5	AAH25578	Nucleotid
C 16	18.4	8.4	38	6	ABK98648	Sequence
C 17	18.4	8.4	38	7	ACA94740	Class C T
C 18	18.4	8.4	38	7	ABX16493	Tet opera
C 19	18.4	8.4	38	8	ACD13899	Tetracycl
C 20	18.4	8.4	38	9	AD52396	Tet opera
C 21	18.4	8.4	46	2	AAH08997	Inulin r
C 22	18.4	8.4	50	4	AAL28949	Human SNP
C 23	18.4	8.4	50	6	ABZ04239	Human leu

C	24	18.4	8.4	50	6	ABZ03712	Human leu
C	25	18.2	8.3	33	9	ADD28164	Reticuloc
C	26	18.2	8.3	41	6	AAL53837	Human cyc
C	27	18.2	8.3	43	7	ABZ21215	Nucleotid
C	28	18.2	8.3	50	6	ABZ04919	Human leu
C	29	18.2	8.3	50	6	ABZ06364	Human leu
C	30	18.2	8.3	50	6	ABZ06168	Human leu
C	31	18	8.2	29	2	AAZ90453	Human sec
C	32	18	8.2	29	4	AAZ59360	Human sec
C	33	18	8.2	29	6	ABA91029	Biotinyla
C	34	18	8.2	42	2	AAV39800	Oligonucl
C	35	18	8.2	50	3	AAZ52575	HCV RNA p
C	36	18	8.2	50	3	AAZ37201	Human PRO
C	37	18	8.2	50	4	AAL34067	Human SNP
C	38	18	8.2	50	4	AAZ54292	Probe #22
C	39	18	8.2	50	8	ACD68329	Novel hum
C	40	18	8.2	50	8	ACH04431	Human sec
C	41	18	8.2	50	8	ACD67975	Novel hum
C	42	18	8.2	50	9	ADC17995	Human PRO
C	43	18	8.2	50	9	ADD70641	Human sec
C	44	18	8.2	50	9	ADD39718	Human sec
C	45	18	8.2	50	9	ADD70164	Human sec

ALIGNMENTS

RESULT 1  
AAL29957/c  
ID AAL29957 standard; DNA; 50 BP.

AC AAL29957;

XX 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #3165.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; Gene therapy; vaccine; amylase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.

OS Homo sapiens.

XX WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US035498.

XX 28-DEC-1999; 99US-0173419P.

XX 27-DEC-2000; 2000US-00173419.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,  
PT autoimmune diseases and infections.

PS Claim 1; Page 2293; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related

CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-protein coupled receptors and thioesterases. The present sequence is one such oligonucleotide. The oligonucleotides and the peptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of the proteins listed above.

CC Disorders that may be prevented, diagnosed and/or treated include multifactorial diseases with a genetic component, such as autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of the bladder, brain, breast, colon and kidney, leukaemia), diseases of the nervous system and an infection of pathogenic organisms

XX

XX Sequence 50 BP; 8 A; 19 C; 11 G; 12 T; 0 U; 0 Other;

Query Match 9.2%; Score 20.2; DB 4; Length 50;  
 Best Local Similarity 75.8%; Pred. No. 6.9e+03;  
 Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 118 GGTGAGTGACATGGAGATACAGCAGCAGT 150  
 |||||  
 Db 39 GGTGGGTGGAATGGCACATGCCAGCAGCACT 7

RESULT 2  
 AAF30355  
 ID AAF30355 standard; DNA; 28 BP.

XX AAF30355;  
 XX  
 DT 14-MAY-2001 (first entry)  
 XX Human checkpoint gene chr 5' PCR primer.

XX Checkpoint with forkhead associated domain and ring finger; Chfr; human;  
 XX mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;  
 XX ubiquitin-protein ligase; PCR primer; ss.

XX Homo sapiens.  
 XX WO200109150-A2.

XX 08-FEB-2001.  
 XX 14-JUN-2000; 2000WO-US016391.

XX 29-JUL-1999; 99US-0146194P.

XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Halazonetis T, Scolnick D;

XX WPI; 2001-182927/18.

XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a  
 XX checkpoint with forkhead-associated domain and ring finger protein, for  
 XX diagnosing tumorigenic cells and in screening for anticancer drugs.

XX Example 3; Page 38; 85pp; English.

XX The present sequence is that of a 5' PCR primer, used with the 3' primer  
 XX given in AAF30356, to amplify a cDNA fragment corresponding to  
 XX nucleotides 352-1055 of the human chr sequence given in AAF30352. The  
 XX chr gene encodes the human mitotic checkpoint protein Chfr (see  
 XX AAB20219), which is required for regulation of the transition of cells  
 XX from prophase to metaphase during mitosis. Loss of expression of Chfr is  
 XX associated with a predisposition to tumorigenesis upon exposure to  
 XX mitotic stress. A set of primers (see AAF30353-76) was used to amplify  
 XX regions spanning the entire chr coding region in order to determine  
 XX whether the chr gene is mutated in any of the human cancer cell lines  
 XX SW480, DLD1, HT29, HCT116, SAOS2, U2OS, IMR5 and NGP. A mutation leading  
 XX to a Val-580 to Met amino acid substitution was identified in the chr  
 XX gene of U2OS cells. Chfr polypeptides and chr nucleic acids are used in

CC methods of diagnosing tumorigenic cells and of screening for drugs which  
 CC can inhibit the activity of Chfr in a cancer cell, rendering it more  
 CC sensitive to additional antitumour therapies

XX Sequence 28 BP; 12 A; 6 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 9.1%; Score 20; DB 5; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 CTGGAAGATACAGCAGCAGCAG 149  
 |||||

Db 9 CTGGAAGATACAGCAGCAGCAG 28

RESULT 3  
 ABL01061  
 ID ABL01061 standard; DNA; 50 BP.

XX ABL01061;

XX 05-MAR-2002 (first entry)

XX Human SNP involving a gap oligonucleotide SEQ ID NO:1052.

XX Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;  
 XX immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;  
 XX autoimmune disease; inflammation; cancer; nervous system disease;  
 XX infection; polymorphic protein; ds.

XX Homo sapiens.

XX WO200138586-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032311.

XX 24-NOV-1999; 99US-0167383P.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI; 2001-355949/37.

XX Isolated human nucleic acids comprising one or more single nucleotide  
 XX polymorphisms, useful for treating a subject suffering from a pathology,  
 XX e.g. autoimmune diseases, ascribed to the presence of a sequence  
 XX polymorphism.

XX Claim 1; Page 568; 674pp; English.

XX ABL00010 to ABL01104 represent human nucleic acid oligonucleotides  
 XX comprising one or more single nucleotide polymorphisms (SNPs). ABB56531  
 XX to ABB56903 represent human peptides encoded by some of the SNP  
 XX oligonucleotides. The sequences from the present invention can have  
 XX immunosuppressive, cytostatic, antiinflammatory, neuroprotective and  
 XX antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides  
 XX and antibodies from the present invention can be used for treating a  
 XX subject suffering from, at risk for, or suspected of, suffering from a  
 XX pathology ascribed to the presence of a sequence polymorphism. The  
 XX pathology may be autoimmune diseases, inflammation, cancer, diseases of  
 XX the nervous system, and infection by pathogenic microorganisms. The SNPs  
 XX are also useful for determining which forms of a characterised  
 XX polymorphism are present in individuals. The antibodies may be used in  
 XX the detection, quantitation and/or cellular or tissue localisation of a  
 XX polymorphic protein (e.g., for use in measuring levels of the polymorphic  
 XX protein within appropriate physiological samples)

XX Sequence 50 BP; 14 A; 3 C; 27 G; 6 T; 0 U; 0 Other;

Query Match 9.0%; Score 19.8; DB 5; Length 50;

```

Best Local Similarity   69.2%; Pred. No. 9.5e+03;
Matches    27; Conservative    0; Mismatches   12; Indels      0; Gaps      0

QY      9  GAGGAAGCGGGAGTGGACCATCGGCGGAGACGAGGTG 47
         |||||
DB       1  GAGGAGGAGGAGTGCGAGGAGGGGGAGAGAGATG 39

RESULT 4
ABV76434/c
ID ID ABV76434 standard; DNA; 41 BP.
XX AC ABV76434;
XX AC
XX DT
XX DE
XX DE
XX XX
XX Human DNA CGG repeat binding protein 10.23 probe, SEQ ID NO: 9.
XX Human; DNA CGG repeat binding protein 10.23; recombinant production;
KW gene therapy; embryonic development disorder; tumour; cancer; cytosolic;
XX probe; ss.
XX OS Homo sapiens.
XX OS
XX PN CN1343717-A.
XX PD
XX PD
XX PF 10-APR-2002.
XX PF 19-SEP-2000; 2000CN-00125256.
XX PF 19-SEP-2000; 2000CN-00125256.
XX PR
XX PR (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX PI Mao Y, Xie Y;
XX DR WPI; 2002-548885/59.
XX DR
XX A novel human DNA CGG-like repeated binder 10.23 polypeptide, useful for
PT treating several diseases e.g. embryo development deformity and tumors.
PF Example 6; Page 22 (Disclosure); 35pp; Chinese.
XX PS
XX CC
XX The invention relates to human DNA CGG repeat binding protein 10.23
CC (ABP58472) and nucleic acids encoding it (ABV76428). The protein has a
CC molecular weight of 10.23 kD. The invention also relates to a method for
CC the recombinant production of the protein, an antagonist of the protein,
CC and the use of the protein, gene and antagonist in therapeutic
CC applications. DNA CGG repeat binding protein 10.23 can be used in the
CC treatment of a variety of diseases such as embryonic development
CC disorders and tumours. Sequences ABV76433-ABV76434 represent human DNA
CC CGG repeat binding protein 10.23 probes used in an exemplification of the
CC invention
XX SQ Sequence 41 BP; 12 A; 11 C; 6 G; 12 T; 0 U; 0 Other;
XX Query Match          8.8%; Score 19.2; DB 6; Length 41;
XX Best Local Similarity 67.5%; Pred. No. 1.4e+04;
XX Matches    27; Conservative    0; Mismatches   13; Indels      0; Gaps      0;

QY      149 GTGGACACTGATTAACTAAGCTGAAGGTGTTGTAAGAAGCA 188
        |||||
DB       40 GTGTAGCATTGGTGTAAGAGGTGCACATTTCCTAAGAAGCA 1

RESULT 5
AAH88323/c
ID ID AAH88323 standard; DNA; 47 BP.
XX AC
XX AC
XX AC
XX AC
XX DT 26-FEB-2002 (first entry)
XX DT

```

[illegible]

PI Bujard H, Gossen M;  
 XX WPI; 1996-087666/09.  
 XX  
 PT New tetracycline-regulated transcription modulators - comprising fusion  
 PT proteins which bind to tet operator sequences to activate or inhibit  
 PT transcription.  
 XX  
 PS Disclosure; Page 76; 112pp; English.  
 XX  
 CC Fusion proteins comprising a first polypeptide which binds to a tet  
 CC operator sequence in the presence of tetracycline or a tetracycline  
 CC analogue, operatively linked to a second polypeptide which either  
 CC activates or inhibits transcription in eukaryotic cells may be used to  
 CC activate or inhibit transcription. Such proteins may be used to regulate  
 CC gene expression in cells and may be particularly useful for gene therapy  
 CC and for expression of gene products in transgenic organisms. Induction of  
 CC gene expression is rapid, efficient and strong, typically 1000-2000 fold.  
 CC The inducing agent does not cause pleiotropic effects or cytotoxicity in  
 CC eukaryotic cells. The sequences of the different classes of tet operator  
 CC sequences are described in AAT11360-64)  
 XX  
 SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;  
 Query Match 8.4%; Score 18.4; DB 2; Length 38;  
 Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 OY 148 AGTGGACAGTGTATTAACAAGCTGAAGGTTGTTAAG 183  
 DB 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3  
 RESULT 9  
 AAT45713/C  
 ID AAT45713 standard; DNA; 38 BP.  
 XX  
 AC AAT45713;  
 DT 18-MAR-1997 (first entry)  
 XX  
 DE Class C Tet operator sequence.  
 XX  
 KW Tet; tetracycline; Tc; operator; transcription; regulation; inducible;  
 KW repressor; gene expression; therapy; transgenic animal; disease model;  
 KW ss.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO9640892-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-US009049.  
 XX  
 PR 07-JUN-1995; 95US-00485971.  
 XX  
 PA (BADI ) BASF AG.  
 XX  
 PI Bujard H, Gossen M, Hillen W, Helbl V, Schnappinger D;  
 XX WPI; 1997-052305/05.  
 XX  
 DR Nucleic acid encoding tetracycline-inducible transcription regulatory  
 PT fusion protein - comprising modified tetracycline repressor able to bind  
 PT mutant tet operator, fused to transcription regulator, useful for  
 PT modulating eukaryotic gene expression.  
 XX  
 PS Disclosure; Page 82; 117pp; English.  
 XX  
 CC AAT45712-T45715 are class A, C, D and E tet operator sequences,  
 CC respectively. The main invention of the specification concerns modified  
 CC Tet repressor (TetR) proteins that bind to modified class B tet operator,  
 CC but also any other class of tet operator may be used. Modified class B  
 CC tet operators were modified at either posn. +4 or +6 by cytosine  
 CC substitutions to give tetO-4C and tetO-6C (see AAT45711 and AAT48478).  
 CC The tet operator sequences were used as targets for modified versions of  
 CC the Tet repressor (TetR). Modified TetR proteins can be fused to a  
 CC transcription regulatory polypeptide and used to control transcription of  
 CC a tetO-4C or tetO-6C linked gene. Nucleic acid encoding a fusion protein,  
 CC as above, may be introduced into a cell and transcription of the protein  
 CC can be controlled by altering the concn. of tetracycline (or an analogue)  
 CC in the cell. This ability to modulate gene expression in a predictable  
 CC way is very useful in gene therapy and for recombinant protein prodn. in  
 CC cultured cells or transgenic animals. The tetracycline (Tc)-inducible  
 CC system is also useful for the prodn. of transgenic animal models for the  
 CC study of disease and also for the study of gene function e.g. during  
 CC differentiation. The Tc-inducible system allows rapid activation of gene  
 CC transcription without cellular toxicity, high concns. of inducer are not  
 CC required  
 XX  
 SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;  
 Query Match 8.4%; Score 18.4; DB 2; Length 38;  
 Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
 Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 OY 148 AGTGGACAGTGTATTAACAAGCTGAAGGTTGTTAAG 183  
 DB 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3  
 RESULT 10  
 AAV60085/C  
 ID AAV60085 standard; DNA; 38 BP.  
 XX  
 AC AAV60085;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 04-DEC-1998 (first entry)  
 XX  
 DE Class C tet operator sequence.  
 XX  
 KW Tet repressor; tetracycline; regulation; expression;  
 KW Tet operator-linked gene; tet operator; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN US5814618-A.  
 XX  
 PD 29-SEP-1998.  
 XX  
 PF 07-JUN-1995; 95US-00485978.  
 XX  
 PR 14-JUN-1993; 93US-00076327.  
 PR 14-JUN-1993; 93US-00076726.  
 PR 14-JUN-1994; 94US-00260452.  
 PR 01-JUL-1994; 94US-00270637.  
 PR 15-JUL-1994; 94US-00275876.  
 PR 03-FEB-1995; 95US-00383754.  
 XX  
 PA (KNOL ) KNOLL AG.  
 PA (BADI ) BASF AG.  
 XX  
 PI Gossen M, Bujard H;  
 XX WPI; 1998-541795/46.  
 XX  
 DR Tetracycline based regulation of gene expression - uses a tetracycline  
 PT operator sequence joined to a gene of interest, the gene of interest  
 PT being induced in the presence, but not absence of the antibiotic.  
 XX  
 PS Disclosure; Fig 5; 63pp; English.  
 XX  
 CC AAV60083-87 represent tet operator sequences of different classes. They  
 CC are used in the course of the invention. The specification describes a



CC method for regulating expression of a Tet (tetracycline) operator-linked  
CC gene in a cell of a subject. The method comprises introducing into the  
CC cell a nucleic acid encoding a fusion protein which inhibits  
CC transcription in eukaryotic cells, the fusion protein comprising a  
CC polypeptide which binds to a Tet operator sequence, operatively linked to  
CC heterologous second polypeptide which inhibits transcription in  
CC eukaryotic cells and modulating the concentration of a tetracycline  
CC (analogue) in the subject. The method is used for the regulation of gene  
CC expression system, using tetracycline (analogues). The system enables a  
CC gene coupled to the system to be induced in the presence of Tet and then  
CC stopped when Tet is removed. (Updated on 25-MAR-2003 to correct PR  
CC field.)  
XX  
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;  
  
Query Match 8.4%; Score 18.4; DB 2; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
  
QY 148 AGTGAACAGTGATTAAACAGCTGAAGGTTGTTAAG 183  
DB 38 AATTAACTGCGATAAACTAGCTTATCGATGATAAG 3  
  
RESULT 11  
AAZ60051/c  
ID AAZ60051 standard; DNA; 38 BP.  
XX  
AC AAZ60051;  
XX  
DT 20-MAR-2003 (revised)  
DT 04-AUG-1999 (first entry)  
XX  
DE Class C tet operator sequence.  
XX  
KW Transgenic mice; transgene; tet operator-linked gene; tetracycline;  
KW mouse-active transcriptional regulatory element; mutant Tet repressor;  
KW gene therapy; genetic disease; acquired disease; cancer; viral disease;  
KW vaccination; rheumatoid arthritis; hypopituitarism; wound healing;  
KW tissue regeneration; cancer; benign prostatic hypertrophy; hemophilia;  
KW erythrocytopenia; artherosclerosis; liver disease; Alzheimer's disease;  
KW Parkinson's disease; human disease model; ds.  
XX  
OS Unidentified.  
XX  
XX US5912411-A.  
XX  
XX 15-JUN-1999.  
XX  
XX 07-JUN-1995; 95US-00487472.  
XX  
XX 14-JUN-1993; 93US-00076327.  
XX 14-JUN-1993; 93US-00076726.  
XX 14-JUN-1994; 94US-00260452.  
XX 01-JUL-1994; 94US-00270637.  
XX 15-JUL-1994; 94US-00275876.  
XX 03-FEB-1995; 95US-00383754.  
XX  
XX (UYHE-) UNIV HEIDELBERG.  
XX  
XX Bujard H, Gossen M;  
XX  
XX WPI; 1999-357232/30.  
XX  
XX Transgenic mice with inducible transgene activity useful for in vitro and  
XX in vivo protein production.  
XX  
XX Disclosure; Fig 5; 63pp; English.  
XX  
XX The specification describes transgenic mice which have a transgene and a  
XX tet operator-linked gene integrated in the genome. The transgene  
XX comprises a mouse-active transcriptional regulatory element linked to a  
XX polynucleotide sequence that encodes a fusion protein which activates

CC transcription of the tet operator-linked gene. The fusion protein  
CC comprises a mutated Tet repressor that binds a tet operator sequence in  
CC the presence of tetracycline (Tc) or a Tc analogue, linked to a  
CC polypeptide that activates transcription in eukaryotic cells. The  
CC transgenic system may be used for gene therapy to treat genes involved in  
CC genetic or acquired diseases. Gene therapy may be used to treat cancer,  
CC viral diseases, for vaccination, and to provide (Tc induced) regulated  
CC doses of a product (e.g. for the treatment or regulation of rheumatoid  
CC arthritis, hypopituitarism, wound healing and tissue regeneration,  
CC cancer, benign prostatic hypertrophy, hemophilia, erythrocytopenia,  
CC atherosclerosis and liver disease, Alzheimer's disease, and Parkinson's  
CC disease). The system may also be used to produce proteins in vivo (e.g.  
CC using mammalian, yeast or fungal cells) or in vitro (e.g. transgenic farm  
CC animals), to produce animal models of human disease, or to produce a  
CC stable cell line for gene cloning. The present sequence is used to  
CC construct the transgenes of the invention. (Updated on 20-MAR-2003 to  
CC correct PF field.)  
XX  
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;  
  
Query Match 8.4%; Score 18.4; DB 2; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
  
QY 148 AGTGAACAGTGATTAAACAGCTGAAGGTTGTTAAG 183  
DB 38 AATTAACTGCGATAAACTAGCTTATCGATGATAAG 3  
  
RESULT 12  
AAZ56132/c  
ID AAZ56132 standard; DNA; 38 BP.  
XX  
AC AAZ56132;  
XX  
DT 27-MAR-2000 (first entry)  
XX  
DE Class C tetracycline operator sequence.  
XX  
KW Class C tetracycline operator; gene expression regulation; cancer;  
KW gene therapy; arthritis; wound healing; tissue regeneration; ds.  
XX  
OS Escherichia coli.  
XX  
XX US6004941-A.  
XX  
XX 21-DEC-1999.  
XX  
XX 07-JUN-1995; 95US-00485740.  
XX  
XX 14-JUN-1993; 93US-00076327.  
XX 14-JUN-1993; 93US-00076726.  
XX 14-JUN-1994; 94US-00260452.  
XX 01-JUL-1994; 94US-00270637.  
XX 15-JUL-1994; 94US-00275876.  
XX 03-FEB-1995; 95US-00383754.  
XX  
XX (BADI ) BASF AG.  
XX  
XX (BADI ) BASF BIORESEARCH CORP.  
XX  
XX (KNOL ) KNOLL AG.  
XX  
XX Gossen M, Bujard H;  
XX  
XX WPI; 2000-085798/07.  
XX  
XX Regulation of gene expression in cells, useful for gene therapy of  
XX diseases, production of proteins in vitro and in vivo and production of  
XX stable cell lines for cloning.  
XX  
XX Disclosure; Fig 5; 64pp; English.  
XX  
XX This is a class C tetracycline operator sequence. The operator sequence  
XX can be used in a method for the regulation of gene expression in a cell,  
CC

CC using tetracycline-regulated fusion proteins. The method involves  
CC obtaining a cell from a subject, and introducing a nucleic acid molecule  
CC into the cell, which operatively links a gene to at least one  
CC tetracycline (tet) operator sequence. A second nucleic acid molecule is  
CC then introduced which encodes a fusion protein, where the fusion protein  
CC comprises a first polypeptide which binds to a tet operator sequence in  
CC the presence of tetracycline, or its analogue, operatively linked to a  
CC second polypeptide which activates transcription in eukaryotic cells to  
CC form a modified cell. The modified cell can then be administered to the  
CC subject, and the concentration of tetracycline (or and analogue) can be  
CC regulated so that the expression of the gene is regulated. The method is  
CC useful for gene therapy of diseases such as cancer and arthritis or for  
CC tissue regeneration and wound healing. The method may also be useful for  
CC the production of proteins in vitro and in vivo and for the production of  
CC stable cell lines for cloning  
XX  
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 3; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAACAAGCTGAAGTTGTTAAG 183  
Db 38 AATTAACTGTGATAAAGCTTATCGATGATAAG 3

RESULT 13  
AAC61972/c  
ID AAC61972 standard; DNA; 38 BP.

AC AAC61972;  
DT 06-MAR-2001 (first entry)

XX Nucleotide sequence of a class C tet operator.

DE Tetracycline; gene therapy; cancer; viral disease; rheumatoid arthritis;  
KW hypopituitarism; wound healing; angiogenesis inhibitor; transgene;  
KW immune response; tet operator; ds.

XX Unidentified.

XX US6136954-A.

XX 24-OCT-2000.

XX 28-SEP-1998; 98US-00162184.

XX 14-JUN-1993; 93US-00076327.

XX 14-JUN-1993; 93US-00076726.

XX 01-JUL-1994; 94US-00260452.

XX 15-JUL-1994; 94US-00270637.

XX 03-FEB-1995; 95US-00383754.

XX 07-JUN-1995; 95US-00485978.

XX (KNOL ) KNOLL AG.

PA (BADI ) BASF AG.

XX Gossen M, Bujard H;

XX WPI; 2001-040240/05.

XX New tetracycline-inducible transcriptional activator fusion proteins  
XX useful for regulating the level of gene expression or in gene therapy for  
XX treating genetic or acquired diseases, e.g. as cancer, viral diseases, or  
XX wound healing.

XX Disclosure; Fig 5; 68pp; English.

XX The specification describes methods for regulating gene expression in  
XX subjects using tetracycline-responsive fusion protein. The fusion protein  
CC

CC comprises a polypeptide that binds to a test operator sequence in the  
CC presence of tetracycline or tetracycline analogues, operatively linked to  
CC a second polypeptide which activates transcription in eukaryotic cells.  
CC The methods are used to turn gene expression on and off, or regulate the  
CC level of gene expression. The system may be used in the study of cellular  
CC development and differentiation in eukaryotic cells, plants and animals,  
CC to regulate expression of site-specific recombinases (e.g. CRE or FLP) to  
CC allow for irreversible modification of the genotype of a transgenic  
CC organism under controlled conditions at a particular stage of  
CC development. The system may further be used in gene therapy, in treating  
CC either genetic or acquired diseases, such as cancer, viral diseases,  
CC rheumatoid arthritis, hypopituitarism, or wound healing, and to  
CC conditionally express suicide gene in cells. The regulatory system is  
CC also used to express angiogenesis inhibitors from within a tumour via a  
CC transgene regulated by this system, and avoid or inhibit an immune  
CC response in subjects receiving treatments. The present sequence  
CC represents a tet operator. It is used in the course of the invention  
XX  
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;

Query Match 8.4%; Score 18.4; DB 4; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAACAAGCTGAAGTTGTTAAG 183  
Db 38 AATTAACTGTGATAAAGCTTATCGATGATAAG 3

RESULT 14

AAH47637/c

ID AAH47637 standard; DNA; 38 BP.

AC AAH47637;

DT 30-NOV-2001 (first entry)

XX Nucleotide sequence of a class C tet operator.

XX Fusion protein; tet operator; tetracycline; gene therapy; anti-cancer;  
XX rheumatoid arthritis; hypopituitarism; wound healing; hemophilia;  
XX diabetes; Alzheimer's disease; tet repressor; ds.

XX Unidentified.

XX US6271348-B1.

XX 07-AUG-2001.

XX 24-JAN-2000; 2000US-00489777.

XX 14-JUN-1993; 93US-00076726.

XX 19-JUN-1993; 93US-00076327.

XX 14-JUN-1994; 94US-00260452.

XX 01-JUL-1994; 94US-00270637.

XX 15-JUL-1994; 94US-00275876.

XX 03-FEB-1995; 95US-00383754.

XX 07-JUN-1995; 95US-00485978.

XX 28-SEP-1998; 98US-00162184.

XX (BADI ) BASF AG.

XX (KNOL ) KNOLL AG.

XX Bujard H, Gossen M;

XX WPI; 2001-556625/62.

XX Fusion protein for inhibiting transcription in eukaryotic cells useful in  
XX gene therapy applications comprises a first polypeptide, which binds to  
XX tet operator sequences, operatively linked to a heterologous second  
XX polypeptide.  
XX Disclosure; Fig 5; 69pp; English.

XX The invention relates to a fusion protein that comprises a first  
CC polypeptide which binds to a tet operator sequence, operatively linked to  
CC a heterologous second polypeptide, which inhibits transcription in  
CC eukaryotic cells. The fusion proteins are tetracycline-responsive and are  
CC useful for regulation of transcription in eukaryotic cells and animals.  
CC The tetracycline (Tc)-controlled regulatory system is useful in various  
CC applications in gene therapy, such as in the treatment of various disease  
CC conditions e.g. rheumatoid arthritis, hypopituitarism, wound healing and  
CC tissue regeneration, anticancer treatments, benign prostatic hyper trophy,  
CC hemophilia, diabetes and arteriosclerosis. They are also useful for bone  
CC marrow support therapy, treatment of central nervous system disorders  
CC e.g. Alzheimer's disease, Parkinson's disease (see AAH47628 for a  
CC detailed description of the uses). The present sequence represents the  
CC nucleotide sequence of a class C tet operator  
XX  
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;  
  
Query Match 8.4%; Score 18.4; DB 4; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
  
QY 148 AGTGGACAGTGATTAACACAGCTGAAGCTTGTAAAG 183  
DB 38 AATTAACTGTGATAAAGCTTATCGATGATAAG 3  
  
RESULT 15  
AAH25578/C  
ID AAH25578 standard; DNA; 38 BP.  
XX  
AC AAH25578;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Nucleotide sequence of a tet operator, class A.  
XX  
KW Transgenic plant; transgene; tet operator-linked gene; Tet repressor;  
KW tetracycline-regulated transcriptional regulatory system; tet operator;  
KW tetracycline; ds.  
XX  
OS Unidentified.  
XX  
PN US6242667-B1.  
XX  
PD 05-JUN-2001.  
XX  
PF 28-SEP-1998; 98US-00161902.  
XX  
PR 14-JUN-1993; 93US-00076327.  
PR 14-JUN-1993; 93US-00076796.  
PR 14-JUN-1994; 94US-00260452.  
PR 01-JUL-1994; 94US-00270637.  
PR 15-JUL-1994; 94US-00275876.  
PR 03-FEB-1995; 95US-00383754.  
PR 07-JUN-1995; 95US-00487472.  
XX  
PA (BADI ) BASF AG.  
PA (KNOL ) KNOLL AG.  
XX  
PI Bujard H, Gossen M;  
XX  
XX WPI; 2001-396837/42.  
DR  
XX  
PT Transgenic plant for regulating the expression of genes, comprises a  
PT transgene integrated into the genome and a tet operator-linked gene in  
PT the genome.  
XX  
XX Disclosure; Fig 5; 65pp; English.  
XX  
XX The specification describes a transgenic plant with a tetracycline-  
CC regulated transcriptional regulatory system. The transgenic plants have a  
CC transgene and a tet operator-linked gene integrated in the genome, which

CC confers a detectable and functional phenotype on the plant when  
CC expressed. The transgene comprises a transcriptional regulatory element  
CC functional in cells of the plant operatively linked to a polynucleotide  
CC sequence encoding a fusion protein that activates transcription of the  
CC tet operator linked gene. The fusion protein comprises a first  
CC polypeptide that is a mutated Tet repressor that binds to a tet operator  
CC sequence in the presence of tetracycline or its analogue, operatively  
CC linked to a second polypeptide that activates transcription in eukaryotic  
CC cells. In the presence of tetracycline, the fusion protein binds to the  
CC tet operator-linked gene and activates transcription of the tet operator-  
CC linked gene such that it is expressed at a level sufficient to confer the  
CC detectable and functional phenotype on the plant. The level of expression  
CC of the tet-operator gene can be downmodulated by depleting tetracycline  
CC from the plant. The transgene can be used to regulate the expression of  
CC genes in the plant. The transgenic plant can be used to analyse the  
CC functions of cellular proteins. The present sequence represents a tet  
CC operator, which can be used to produce transgenic plants of the invention  
XX  
SQ Sequence 38 BP; 13 A; 6 C; 5 G; 14 T; 0 U; 0 Other;  
  
Query Match 8.4%; Score 18.4; DB 5; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.6e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
  
QY 148 AGTGGACAGTGATTAACACAGCTGAAGCTTGTAAAG 183  
DB 38 AATTAACTGTGATAAAGCTTATCGATGATAAG 3  
  
Search completed: May 26, 2004, 05:53:43  
Job time : 186.951 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	20	9.1	47	4	US-09-641-638-63	Sequence 687, App	
C 2	18.8	8.6	38	1	US-08-117-361C-23	Sequence 23, Appl	
	18.4	8.4	38	1	US-08-485-971-13	Sequence 13, Appl	
	18.4	8.4	38	1	US-08-275-876-13	Sequence 13, Appl	
C 4	18.4	8.4	38	1	US-08-383-754-13	Sequence 13, Appl	
C 5	18.4	8.4	38	1	US-08-485-978-13	Sequence 13, Appl	
C 6	18.4	8.4	38	2	US-08-486-814-13	Sequence 13, Appl	
C 7	18.4	8.4	38	2	US-08-487-472-13	Sequence 13, Appl	
C 8	18.4	8.4	38	3	US-08-485-740-13	Sequence 13, Appl	
C 9	18.4	8.4	38	3	US-08-485-476-13	Sequence 13, Appl	
C 10	18.4	8.4	38	3	US-09-162-184-13	Sequence 13, Appl	
C 11	18.4	8.4	38	3	US-09-161-902-13	Sequence 13, Appl	
C 12	18.4	8.4	38	3	US-09-489-777A-13	Sequence 13, Appl	
C 13	18.4	8.4	38	5	PCN-US95-08179-13	Sequence 13, Appl	
C 14	18.4	8.4	46	1	US-08-242-409-12	Sequence 12, Appl	
C 15	18.4	8.4	46	5	PCN-US95-US835-12	Sequence 12, Appl	
C 16	18.4	8.4	47	4	US-09-641-638-1120	Sequence 1120, Ap	
C 17	18.4	8.4	47	4	US-09-422-978-1696	Sequence 1696, Ap	
C 18	17.8	8.1	49	1	US-08-171-389-54	Sequence 54, Appl	
C 19	17.8	8.1	49	1	US-08-123-936-54	Sequence 54, Appl	
C 20	17.8	8.1	49	2	US-08-475-228A-54	Sequence 54, Appl	
C 21	17.8	8.1	49	3	US-08-482-080A-54	Sequence 54, Appl	
C 22	17.8	8.1	49	4	US-09-354-947-54	Sequence 54, Appl	
C 23	17.8	8.1	49	5	PCN-US93-12388-54	Sequence 54, Appl	
C 24	17.4	7.9	47	4	US-09-671-317-553	Sequence 553, App	
C 25	17.4	7.9	50	1	US-08-171-389-546	Sequence 546, App	
C 26	17.4	7.9	50	2	US-08-123-936-546	Sequence 546, App	
C 27	17.4	7.9	50	2	US-08-475-228A-546	Sequence 546, App	

;; TITLE OF INVENTION: Endothelin B1 (ETb1) Receptor Polypeptide  
;; TITLE OF INVENTION: Compositions, Methods, and Uses Thereof  
;; NUMBER OF SEQUENCES: 27  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Chiron Corporation  
;; STREET: 4560 Horton Street  
;; CITY: Emeryville  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 946087  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25 &  
;; SOFTWARE: Word 5.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/117,361C  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Chung, Ling-Fong  
;; REGISTRATION NUMBER: 36,482  
;; REFERENCE/DOCKET NUMBER: 0945.001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (510) 923-2704  
;; TELEFAX: (510) 655-3542  
;; INFORMATION FOR SEQ ID NO: 23:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 38 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; US-08-117-361C-23

Query Match 8.6%; Score 18.8; DB 1; Length 38;  
Best Local Similarity 68.4%; Pred. No. 1.7e+03;  
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 65 GCATATAACTGGTCTCTGGAGATCACTGTAGAAATGTA 102  
DB 38 GCACAGCTGCATATCTGACGATTAATAAGAAATGTA 1

RESULT 3  
US-08-485-971-13/c  
; Sequence 13, Application US/08485971  
; Patent No. 5589362  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; APPLICANT: Gossen, Manfred  
; APPLICANT: Hillen, Wolfgang  
; APPLICANT: Helbl, Vera  
; APPLICANT: Schnappinger, Dirk  
; TITLE OF INVENTION: Tetracycline-Regulated Transcriptional  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,971  
; FILING DATE:  
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/383,754  
;; FILING DATE: 03-FEB-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/270,637  
;; FILING DATE: 01-JULY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/260,452  
;; FILING DATE: 14-JUNE-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/076,327  
;; FILING DATE: 14-JUNE-1993  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/076,726  
;; FILING DATE: 14-JUNE-1993  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: DeConti, Giulio A. Jr.  
;; REGISTRATION NUMBER: 31,503  
;; REFERENCE/DOCKET NUMBER: BBI-009CP7  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)227-7400  
;; TELEFAX: (617)227-5941  
;; INFORMATION FOR SEQ ID NO: 13:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 38 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA  
;; US-08-485-971-13

Query Match 8.4%; Score 18.4; DB 1; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.4e+03;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 ACTGGAACAGTGATTAACAAGCTGAAGGTGTTAAG 183  
DB 38 AATTTAACTGTGATAAAGTATGCTTATCGATGATGAAG 3

RESULT 4  
US-08-275-876-13/c  
; Sequence 13, Application US/08275876  
; Patent No. 5654168  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; APPLICANT: Gossen, Manfred  
; TITLE OF INVENTION: Tetracycline-Inducible Transcriptional  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/275,876  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US  
; FILING DATE: 01-JULY-94

```
/ CLASSIFICATION: 436
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DeConti, Giulio A. Jr.
/ REGISTRATION NUMBER: 31,503
/ REFERENCE/DOCKET NUMBER: BBI-009CP
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)227-5941
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 38 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-08-275-876-13

Query Match      8.4%; Score 18.4; DB 1; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.4e+03;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAAACAAGCTGAAGCTGTGTTAAG 183
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 5
US-08-383-754-13/c
/ Sequence 13, Application US/08383754
/ Patent No. 5789156
/ GENERAL INFORMATION:
/ APPLICANT: Bujard, Hermann
/ APPLICANT: Gossen, Manfred
/ TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
/ NUMBER OF SEQUENCES: 28
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 60 State Street, suite 510
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/383,754
/ FILING DATE:
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/275,876
/ FILING DATE: 15-JULY-94
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/270,637
/ FILING DATE: 01-JULY-94
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/260,452
/ FILING DATE: 14-JUNE-94
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/076,327
/ FILING DATE: 14-JUNE-93
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/076,726
/ FILING DATE: 14-JUNE-93
/ CLASSIFICATION: 436
/ ATTORNEY/AGENT INFORMATION:

NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-383-754-13

Query Match      8.4%; Score 18.4; DB 1; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.4e+03;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAAACAAGCTGAAGCTGTGTTAAG 183
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 6
US-08-485-978-13/c
/ Sequence 13, Application US/08485978
/ Patent No. 5814618
/ GENERAL INFORMATION:
/ APPLICANT: Bujard, Hermann
/ APPLICANT: Gossen, Manfred
/ TITLE OF INVENTION: Methods for Regulating Gene Expression
/ NUMBER OF SEQUENCES: 28
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 60 State Street, Suite 510
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/485,978
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/383,754
/ FILING DATE: 03-FEB-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/275,876
/ FILING DATE: 15-JULY-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/270,637
/ FILING DATE: 01-JULY-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/260,452
/ FILING DATE: 14-JUNE-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/076,327
/ FILING DATE: 14-JUNE-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/076,726
/ FILING DATE: 14-JUNE-1993
/ CLASSIFICATION: 435
```

ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-009CP6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-485-978-13

Query Match 8.4%; Score 18.4; DB 1; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.4e+03;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAACAAGCTGAAGTTGTTAAG 183  
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 7  
US-08-486-814-13/c  
Sequence 13, Application US/08486814  
Patent No. 5866755  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
APPLICANT: Gossen, Manfred  
TITLE OF INVENTION: Animals Transgenic for a Tetracycline-  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486.814  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/383,754  
FILING DATE: 03-FEB-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/275,876  
FILING DATE: 15-JULY-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/270,637  
FILING DATE: 01-JULY-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,452  
FILING DATE: 14-JUNE-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076,327  
FILING DATE: 14-JUNE-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076,726  
FILING DATE: 14-JUNE-1993

Controlled Transcr

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-009CP4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-486-814-13

Query Match 8.4%; Score 18.4; DB 2; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.4e+03;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAACAAGCTGAAGTTGTTAAG 183  
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 8  
US-08-487-472-13/c  
Sequence 13, Application US/08487472  
Patent No. 5912411  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
APPLICANT: Gossen, Manfred  
TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible Transcription  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,472  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/383,754  
FILING DATE: 03-FEB-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/275,876  
FILING DATE: 15-JULY-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/270,637  
FILING DATE: 01-JULY-94  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,452  
FILING DATE: 14-JUNE-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076,327  
FILING DATE: 14-JUNE-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076,726  
FILING DATE: 14-JUNE-1993

```
/
/ FILING DATE: 14-JUNE-1993
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DeConti, Giulio A. Jr.
/ REGISTRATION NUMBER: 31,503
/ REFERENCE/DOCKET NUMBER: BBI-009CP3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)227-5941
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 38 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ US-08-487-472-13

Query Match      8.4%; Score 18.4; DB 2; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.4e+03;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGAAACAGTGATTAAACAAGCTGAAGGTTGTTAAG 183
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 9
US-08-485-740-13/c
/ Sequence 13, Application US/08485740
/ Patent No. 6004941
/ GENERAL INFORMATION:
/ APPLICANT: Bujard, Hermann
/ APPLICANT: Gossen, Manfred
/ TITLE OF INVENTION: Methods for Regulating Gene Expression
/ NUMBER OF SEQUENCES: 28
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 60 State Street, Suite 510
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/485,740
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/383,754
/ FILING DATE: 03-FEB-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/275,876
/ FILING DATE: 15-JULY-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/270,637
/ FILING DATE: 01-JULY-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/260,452
/ FILING DATE: 14-JUNE-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/076,327
/ FILING DATE: 14-JUNE-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/076,726
/ FILING DATE: 14-JUNE-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DeConti, Giulio A. Jr.
/ REGISTRATION NUMBER: 31,503
/ REFERENCE/DOCKET NUMBER: BBI-009C6CN
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617)227-7400

/
/ APPLICATION NUMBER: US 08/076,726
/ FILING DATE: 14-JUNE-1993
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DeConti, Giulio A. Jr.
/ REGISTRATION NUMBER: 31,503
/ REFERENCE/DOCKET NUMBER: BBI-009CPS
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)227-5941
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 38 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ US-08-485-740-13

Query Match      8.4%; Score 18.4; DB 3; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.4e+03;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGAAACAGTGATTAAACAAGCTGAAGGTTGTTAAG 183
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 10
US-09-162-184-13/c
/ Sequence 13, Application US/09162184A
/ Patent No. 6136954
/ GENERAL INFORMATION:
/ APPLICANT: Bujard, Hermann
/ APPLICANT: Gossen, Manfred
/ TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
/ NUMBER OF SEQUENCES: 37
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 28 State Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/162,184A
/ FILING DATE: 28-Sep-1998
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/383,754
/ FILING DATE: 03-FEB-1995
/ APPLICATION NUMBER: US 08/275,876
/ FILING DATE: 15-JULY-1994
/ APPLICATION NUMBER: US 08/270,637
/ FILING DATE: 01-JULY-1994
/ APPLICATION NUMBER: US 08/260,452
/ FILING DATE: 14-JUNE-1994
/ APPLICATION NUMBER: US 08/076,327
/ FILING DATE: 14-JUNE-1993
/ APPLICATION NUMBER: US 08/076,726
/ FILING DATE: 14-JUNE-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: DeConti, Giulio A. Jr.
/ REGISTRATION NUMBER: 31,503
/ REFERENCE/DOCKET NUMBER: BBI-009C6CN
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617)227-7400
```



TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-162-184-13

Query Match 8.4%; Score 18.4; DB 3; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.4e+03;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAAACAGCTGAAGGTTGTTAAG 183  
DB 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 11  
US-09-161-902-13/c  
Sequence 13, Application US/09161902  
Patent No. 6242867  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
Gossen, Manfred  
TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible Transcription  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/161,902  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,472  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/275,876  
FILING DATE: 15-JULY-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/270,637  
FILING DATE: 01-JULY-94  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,452  
FILING DATE: 14-JUNE-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076,327  
FILING DATE: 14-JUNE-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076,726  
FILING DATE: 14-JUNE-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
FILING DATE: 14-JUNE-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-009CP3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-161-902-13

Query Match 8.4%; Score 18.4; DB 3; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.4e+03;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGATTAAACAGCTGAAGGTTGTTAAG 183  
DB 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 12  
US-09-489-777A-13/c  
Sequence 13, Application US/09489777A  
Patent No. 6271348  
GENERAL INFORMATION:  
APPLICANT: Bujard, Hermann  
Gossen, Manfred  
TITLE OF INVENTION: Tetracycline-Inducible Transcriptional  
Inhibitor Fusion Proteins  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/489,777A  
FILING DATE: 24-Jan-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/162,184  
FILING DATE: 28-SEP-1998  
APPLICATION NUMBER: US 08/485,978  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/383,754  
FILING DATE: 03-FEB-1995  
APPLICATION NUMBER: US 08/275,876  
FILING DATE: 15-JULY-1994  
APPLICATION NUMBER: US 08/270,637  
FILING DATE: 01-JULY-1994  
APPLICATION NUMBER: US 08/260,452  
FILING DATE: 14-JUNE-1994  
APPLICATION NUMBER: US 08/076,327  
FILING DATE: 14-JUNE-1993  
APPLICATION NUMBER: US 08/076,726  
FILING DATE: 14-JUNE-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-009C6CNDV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:

```

,
, LENGTH: 38 base pairs
,
, TYPE: nucleic acid
,
, STRANDEDNESS: double
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: DNA
,
, SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-489-777A-13

```

Query Match 8.4%; Score 18.4; DB 3; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.4e+03;  
Matches 25; Conservative 0; Mismatches 11. Indels

QY 148 AGTGGACAGTGATTACAAGCTGAAGTTGTTAAG 183  
| | | | | | | | | | | | | | | | | | | | | |  
Db 38 AATTAACTGTGTAATACTAGCTTATCGATGATGAAG 3  
| | | | | | | | | | | | | | | | | | | | | |

RESULT 13  
PCT-US95-08179-13/c  
; Sequence 13, Application PC/TUS9508179  
; GENERAL INFORMATION:  
; APPLICANT:

AFFILIANT: Tetracycline-Regulated Transcriptional  
 TITLE OF INVENTION: Modulators  
 NUMBER OF SEQUENCES: 28  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street, suite 510  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA

```

? ZIP: 02109-1875
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII Text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/08179
? FILING DATE:
?
```

/ FILING DATE:  
 / CLASSIFICATION:  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: To be assigned  
 / FILING DATE: 07-JUN-1995  
 / CLASSIFICATION:  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: US 08/383,754  
 / FILING DATE: 03-FEB-1995

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/275,876  
FILING DATE: 15-JULY-94

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/270,637  
FILING DATE: 01-JUN-94

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-009C2PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 38 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA

```

PCT-US95-08179-13

```

; Length 38;
;
; Indels 0; Gaps 0;

```

3  
183

transcriptional

	Query Match	8.4%	Score 18.4;	DB 5	Length 38;
	Best Local Similarity	69.4%	Pred No. 2.4e+03;		
	Matches 25;	Conservative	0;	Mismatches 11;	Indels
QY	148	AGTGGACAGGATTAACAGCTGACGGTGTGTTAA	183		
	↑		↑		
	↑		↑		
Db	38	AAATTAACTGTGATAAACTAGCTTATCGATGATAA	3		

RESULT 14  
US-08-242-409-12

Patent No. 5496831  
GENERAL INFORMATION:  
APPLICANT: Alexander-Bridges, Maria C.

APPLICANT: Zhao, Hui-Fen  
TITLE OF INVENTION: INHIBITION OF INSULIN-INDUCED  
ADIPOSIS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street

CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/242,409  
 FILING DATE:  
 CLASSIFICATION: 514  
 PRICE APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Clark, Paul T.

```
,
, REGISTRATION NUMBER: 30,162
, REFERENCE/DOCKET NUMBER: 00786/238001
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (617) 542-5070
, TELEFAX: (617) 542-8906
, TELEX: 200154
, INFORMATION FOR SEQ ID NO: 12:
, SEQUENCE CHARACTERISTICS:
,
```

```

;
; LENGTH: 46
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;

```

US-08-242-409-12

Query Match	8.4%;	Score 18.4;	DB 1;	Length 46
Best Local Similarity	63.6%;	Pred. No. 2.7e+03;		

	Matches	28; Conservative	0; Mismatches	16; Indels
Qy	20	AGTGGACATCGGGGGAGACAGAGTTGCGACCTTTCTTTCCCC	63	
Db	2	AGGGGTTGAAGGGGGAGAGTCGAAACTTTCTTTCTTTCCCC	45	
RESULT 15				

PCT-US95-05835-12  
; Sequence 12, Application PC/TUS9505835

```

: GENERAL INFORMATION:
:
: APPLICANT: Alexander-Bridges, Maria C.
:
: APPLICANT: Zhao, Hui-Fen
: TITLE OF INVENTION: INHIBITION OF INSULIN-
: TITLE OF INVENTION: INDUCED ADIPOSIS
:
:

```

; NUMBER OF SEQUENCES: 12

Search completed: May 26, 2004, 09:20:40  
Job time : 43.3039 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:41:55 ; Search time 955.965 Seconds  
(without alignments)  
1041.142 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_181\_399

Perfect score: 219

Sequence: 1 gtctctgtgagagcggga.....tacagactgggatgtcatc 219

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 1612472

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	9.1	47	16	US-10-170-097-687
2	18.6	8.5	40	12	US-10-035-833A-3666
3	18.4	8.4	38	9	US-09-874-389-13
4	18.4	8.4	38	10	US-09-921-650-13
5	18.4	8.4	38	11	US-09-241-347-13
6	18.4	8.4	38	15	US-10-032-393-64
7	18.4	8.4	38	15	US-10-301-515-10
8	18.4	8.4	38	15	US-10-326-671-53
9	18.4	8.4	47	16	US-10-170-097-1120
10	18.4	8.4	47	16	US-10-349-143-1696
11	18.4	8.4	50	16	US-10-131-827-3703
12	18.4	8.4	50	16	US-10-131-827-4230
13	18.2	8.3	46	13	US-10-252-155-739
14	18.2	8.3	46	13	US-10-252-155-740

15 18.2 8.3 50 16 US-10-131-827-4910 Sequence 4910, Ap  
16 18.2 8.3 50 16 US-10-131-827-6159 Sequence 6159, Ap  
17 18.2 8.3 50 16 US-10-131-827-6355 Sequence 6355, Ap  
18 18 8.2 29 9 US-09-729-674-247 Sequence 247, Ap  
19 18 8.2 42 9 US-09-179-536B-318 Sequence 318, Ap  
20 18 8.2 42 10 US-09-297-576A-318 Sequence 318, Ap  
21 18 8.2 50 10 US-09-946-374-126 Sequence 126, Ap  
22 18 8.2 50 12 US-10-015-395A-126 Sequence 126, Ap  
23 18 8.2 50 13 US-10-006-485A-126 Sequence 126, Ap  
24 18 8.2 50 13 US-10-013-907A-126 Sequence 126, Ap  
25 18 8.2 50 13 US-10-015-499A-126 Sequence 126, Ap  
26 18 8.2 50 13 US-10-226-254A-126 Sequence 126, Ap  
27 18 8.2 50 15 US-10-006-856A-126 Sequence 126, Ap  
28 18 8.2 50 15 US-10-006-818A-126 Sequence 126, Ap  
29 18 8.2 50 15 US-10-013-393A-126 Sequence 126, Ap  
30 18 8.2 50 15 US-10-015-869A-126 Sequence 126, Ap  
31 18 8.2 50 15 US-10-012-121A-126 Sequence 126, Ap  
32 18 8.2 50 15 US-10-006-116A-126 Sequence 126, Ap  
33 18 8.2 50 15 US-10-006-117A-126 Sequence 126, Ap  
34 18 8.2 50 15 US-10-017-527A-126 Sequence 126, Ap  
35 18 8.2 50 15 US-10-013-913A-126 Sequence 126, Ap  
36 18 8.2 50 15 US-10-007-194A-126 Sequence 126, Ap  
37 18 8.2 50 15 US-10-013-430A-126 Sequence 126, Ap  
38 18 8.2 50 15 US-10-011-671A-126 Sequence 126, Ap  
39 18 8.2 50 15 US-10-012-755A-126 Sequence 126, Ap  
40 18 8.2 50 15 US-10-015-386A-126 Sequence 126, Ap  
41 18 8.2 50 15 US-10-011-692A-126 Sequence 126, Ap  
42 18 8.2 50 15 US-10-006-768A-126 Sequence 126, Ap  
43 18 8.2 50 15 US-10-017-610A-126 Sequence 126, Ap  
44 18 8.2 50 15 US-10-006-063A-126 Sequence 126, Ap  
45 18 8.2 50 15 US-10-020-063A-126 Sequence 126, Ap

#### ALIGNMENTS

#### RESULT 1

US-10-170-097-687  
; Sequence 687, Application US/10170097  
; Publication No. US20030228582A1  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Chumakov, Ilva  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: ALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING  
; FILE REFERENCE: GEN-T114XC2D1  
; CURRENT APPLICATION NUMBER: US/10170,097  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US 09/641,638  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: US 09/502,330  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: US 60/133,200  
; PRIOR FILING DATE: 1999-05-07  
; PRIOR APPLICATION NUMBER: US 09/275,267  
; PRIOR FILING DATE: 1999-03-23  
; PRIOR APPLICATION NUMBER: US 60/119,917  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 1304  
; SOFTWARE: Patent.pm  
; SEQ ID NO 687  
; LENGTH: 47  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 24  
; OTHER INFORMATION: 10-347-271 : polymorphic base A or T  
US-10-170-097-687

Query Match 9.1%; Score 20; DB 16; Length 47;

Best Local Similarity 63.0%; Pred. No. 8.5e+03;  
Matches 29; Conservative 1; Mismatches 16; Indels 0; Gaps 0;  
QY 50 ACCTTCCTCCCGGCAATAAAGTCTCTGGAGATCACTGTAG 95  
|||||  
Db 2 ACCTTCCTCCCGGCTGGCAGTCTCTCTCTAGTAGTGG 47  
|||||  
RESULT 2  
US-10-035-833A-3666  
; Sequence 3666, Application US/10035833A  
; Publication No. US20040072156A1  
; GENERAL INFORMATION:  
; APPLICANT: Nakamura, Yuho  
; APPLICANT: Sekine, Akihiro  
; APPLICANT: Iida, Aritoshi  
; APPLICANT: Saito, Osamu  
; TITLE OF INVENTION: Detection of Genetic Polymorphisms  
; FILE REFERENCE: FORS-06904  
; CURRENT APPLICATION NUMBER: US/10/035,833A  
; CURRENT FILING DATE: 2001-12-27  
; NUMBER OF SEQ ID NOS: 7669  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3666  
; LENGTH: 40  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (21)..(21)  
; OTHER INFORMATION: g is present or absent.  
US-10-035-833A-3666  
Query Match 8.5%; Score 18.6; DB 12; Length 40;  
Best Local Similarity 72.7%; Pred. No. 2.4e+04;  
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 7 CTGAGAGCGGAGTGACCATCGGCGGAGA 39  
|||||  
Db 7 CAGAGTAGGGGGGTGGAGGTGGGAGGGA 39  
|||||  
RESULT 3  
US-09-874-389-13/c  
; Sequence 13, Application US/09874389  
; Patent No. US20020152489A1  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; Gossen, Manfred  
; TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible  
; Transcription  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/874,389  
; FILING DATE: 26-Dec-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/161,902  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 08/275,876  
; FILING DATE: 15-JULY-1994

APPLICATION NUMBER: US 08/270,637  
FILING DATE: 01-JULY-94  
APPLICATION NUMBER: US 08/260,452  
FILING DATE: 14-JUNE-1994  
APPLICATION NUMBER: US 08/076,327  
FILING DATE: 14-JUNE-1993  
APPLICATION NUMBER: US 08/076,726  
FILING DATE: 14-JUNE-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: DeConti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: BBI-009CP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-874-389-13  
Query Match 8.4%; Score 18.4; DB 9; Length 38;  
Best Local Similarity 69.4%; Pred. No. 2.8e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 148 AGTGAACAGTGATTAACAAGCTGAAGTGTGTTAAG 183  
|||||  
Db 38 AATTACTGTGATAAAGTACTGATGATGAAG 3  
|||||  
RESULT 4  
US-09-921-650-13/c  
; Sequence 13, Application US/09921650  
; Publication No. US20030022315A1  
; GENERAL INFORMATION:  
; APPLICANT: Bujard, Hermann  
; Gossen, Manfred  
; TITLE OF INVENTION: Tetracycline-Inducible Transcriptional  
; Activator Fusion Proteins  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/921,650  
; FILING DATE: 03-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/912,650  
; FILING DATE: 2001-08-03  
; APPLICATION NUMBER: US 08/485,978  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/383,754  
; FILING DATE: 03-FEB-1995  
; APPLICATION NUMBER: US 08/275,876  
; FILING DATE: 15-JULY-1994  
; APPLICATION NUMBER: US 08/270,637  
; FILING DATE: 01-JULY-1994  
; APPLICATION NUMBER: US 08/260,452  
; FILING DATE: 14-JUNE-1994

```

APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009C6NDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-921-650-13

Query Match      8.4%; Score 18.4; DB 10; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGCTGATTAACAAGCTGAAGGTGTAAAG 183
DB 38 AATTAACTGTGATAAAGTAGCTTATCGATGATAAG 3

RESULT 5
US-09-241-347-13/c
Sequence 13, Application US/09241347
Publication No. US20040003417A1
GENERAL INFORMATION:
APPLICANT: Sujard, Hermann
Gossen, Manfred
TITLE OF INVENTION: Animals Transgenic for a Tetracycline-
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,347
FILING DATE: 02-Feb-1999
CLASSIFICATION: 800
PRIORITY INFORMATION:
APPLICATION NUMBER: US/08/486,814
FILING DATE: UNKNOWN
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503

```

---

```

APPLICATION NUMBER: BBI-009CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-241-347-13

Query Match      8.4%; Score 18.4; DB 11; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGCTGATTAACAAGCTGAAGGTGTAAAG 183
DB 38 AATTAACTGTGATAAAGTAGCTTATCGATGATAAG 3

RESULT 6
US-10-032-393-64/c
Sequence 64, Application US/10032393
Publication No. US20030027286A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
Wall, Daniel
APPLICANT: Gross, Molly
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
FILE REFERENCE: ELITRA.010A
CURRENT APPLICATION NUMBER: US/10/032,393
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/259,434
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: 09/948,993
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 68
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 64
LENGTH: 38
TYPE: DNA
ORGANISM: Plasmid pSC101
US-10-032-393-64

Query Match      8.4%; Score 18.4; DB 15; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGCTGATTAACAAGCTGAAGGTGTAAAG 183
DB 38 AATTAACTGTGATAAAGTAGCTTATCGATGATAAG 3

RESULT 7
US-10-301-516-10/c
Sequence 10, Application US/10301516
Publication No. US20030180756A1
GENERAL INFORMATION:
APPLICANT: SHI, YANG
APPLICANT: SUI, GUANGCHAO
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SUPPRESSING EUKARYOTIC GENE
FILE REFERENCE: HMV-084.01
CURRENT APPLICATION NUMBER: US/10/301,516
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: 60/366,478
PRIOR FILING DATE: 2002-03-21
NUMBER OF SEQ ID NOS: 39

```

```
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic tet
; OTHER INFORMATION: operator sequence
US-10-301-516-10

Query Match      8.4%; Score 18.4; DB 15; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGTATTACACAGCTGAAGGTTGTTAAG 183
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 8
US-10-326-671-53/c
; Sequence 53, Application US/10326671
; Publication No. US20030186281A1
; GENERAL INFORMATION:
; APPLICANT: Hilten, Wolfgang
; TITLE OF INVENTION: MODIFIED TETRACYCLINE REPRESSOR PROTEIN COMPOSITIONS AND METHODS
; FILE REFERENCE: 10182-022-999
; CURRENT APPLICATION NUMBER: US/10/326,671
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/343,278
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 459
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 53
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Plasmid pSC101
US-10-326-671-53

Query Match      8.4%; Score 18.4; DB 15; Length 38;
Best Local Similarity 69.4%; Pred. No. 2.8e+04;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 148 AGTGGACAGTGTATTACACAGCTGAAGGTTGTTAAG 183
Db 38 AATTAACTGTGATAAACTAGCTTATCGATGATAAG 3

RESULT 9
US-10-170-097-1120
; Sequence 1120, Application US/10170097
; Publication No. US20030228582A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
```

```
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 1120
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 12-44-67 : polymorphic base T or C
US-10-170-097-1120

Query Match      8.4%; Score 18.4; DB 16; Length 47;
Best Local Similarity 73.3%; Pred. No. 3e+04;
Matches 22; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 63 CAGCAATAAACTGCTCTCTGGAGATCACTG 92
Db 14 CAGCCAGAAGYTGCTGTGTGGATGCGACTG 43

RESULT 10
US-10-349-143-1696
; Sequence 1696, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020Cp1
; CURRENT APPLICATION NUMBER: US/10/349,143
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US/09/422,978
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1696
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-5709-80 : polymorphic base A or G
US-10-349-143-1696

Query Match      8.4%; Score 18.4; DB 16; Length 47;
Best Local Similarity 60.9%; Pred. No. 3e+04;
Matches 28; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

QY 140 CCAGCACCAGTGGAAACAGTGTAACTCAAGCTGAAGGTTGTTAAGAA 185
Db 1 CTAGCAGAACTGGATCATGATCTACAAAATGAAGAGGCTTTGCA 46

RESULT 11
US-10-131-827-3703/c
; Sequence 3703, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; FILE REFERENCE: 506612000120
```

; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3703  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-3703

Query Match 8.4%; Score 18.4; DB 16; Length 50;  
Best Local Similarity 63.6%; Pred. No. 3.1e+04;  
Matches 28; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 37 AGACGAGGTTCCGACCTTCTCCCTCCCGACGAATAAAGTGTCTC 80  
DB 47 AGACGAGTTCGAGAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 4

## RESULT 12

US-10-131-827-4230/c  
; Sequence 4230, Application US/10131827  
; Publication No. US20040009479A1  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; FILE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4230  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-4230

Query Match 8.4%; Score 18.4; DB 16; Length 50;  
Best Local Similarity 69.4%; Pred. No. 3.1e+04;  
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 50 ACCTTCTCTCCCGACGAATAAAGTGTCTCTGAG 85  
DB 46 ACATTTCATCCCCAGTAATGAAGAGACTCAAGAG 11

## RESULT 13

US-10-252-155-739  
; Sequence 739, Application US/10252155  
; Publication No. US20040068096A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT  
; FILE OF INVENTION: MULTI-DRUG RESISTANT PROTEINS  
; FILE REFERENCE: D0152 NP  
; CURRENT APPLICATION NUMBER: US/10/252,155  
; CURRENT FILING DATE: 2002-09-20  
; PRIOR APPLICATION NUMBER: US 60/324,172  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: US 60/333,700  
; PRIOR FILING DATE: 2001-11-27

; NUMBER OF SEQ ID NOS: 783  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 739  
; LENGTH: 46  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-252-155-739

Query Match 8.3%; Score 18.2; DB 13; Length 46;  
Best Local Similarity 66.7%; Pred. No. 3.5e+04;  
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 181 AAGAAGCAGACATGCCCTTTACAGACTGGGATGTCATC 219  
DB 2 AGGAACAGCTATGACCTTGACATACATTGTGTTTCATC 40

## RESULT 14

US-10-252-155-740  
; Sequence 740, Application US/10252155  
; Publication No. US20040068096A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT  
; FILE OF INVENTION: MULTI-DRUG RESISTANT PROTEINS  
; FILE REFERENCE: D0152 NP  
; CURRENT APPLICATION NUMBER: US/10/252,155  
; CURRENT FILING DATE: 2002-09-20  
; PRIOR APPLICATION NUMBER: US 60/324,172  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: US 60/333,700  
; PRIOR FILING DATE: 2001-11-27  
; NUMBER OF SEQ ID NOS: 783  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 740  
; LENGTH: 46  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-252-155-740

Query Match 8.3%; Score 18.2; DB 13; Length 46;  
Best Local Similarity 66.7%; Pred. No. 3.5e+04;  
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 181 AAGAAGCAGACATGCCCTTTACAGACTGGGATGTCATC 219  
DB 2 AGGAACAGCTATGACCTTGACATACATTGTGTTTCATC 40

## RESULT 15

US-10-131-827-4910  
; Sequence 4910, Application US/10131827  
; Publication No. US20040009479A1  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; FILE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; CURRENT FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4910  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-4910



US-10-131-827-4910

Query Match 8.3%; Score 18.2; DB 16; Length 50;  
 Best Local Similarity 61.7%; Pred. No. 3.6e+04;  
 Matches 29; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qv 91 TGTAGAAATTGTGTGGATGAAATCAAGTCAGGTGACACTGGAAGA 137  
 |||||  
 Db 2 TGGAGCGTTTTTGTGTTGAGATATTAGCTCAGGTCAATTCCAAGA 48  
 |||||

Search completed: May 26, 2004, 10:22:39  
 Job time : 955.965 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:12:53 ; Search time 1520.36 Seconds  
(without alignments)  
4301.492 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_181\_399

Perfect score: 219  
Sequence: 1 gtctctctgaggaagcggga.....tacagactgggagtgatc 219

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estab:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	10.0	50	28	AZ481318 1M0303B05
2	21.6	9.9	50	9	AUI02838
C 3	20.8	9.5	43	9	AIZ66869 uJ08a10.x
C 4	20.8	9.5	50	9	AUI07911 AUI07911

C	5	20.2	9.2	50	9	AUI05891
	6	19.4	8.9	50	9	AUI02843
	7	19.4	8.9	50	9	AUI02844
	8	18.8	8.6	40	9	AA137304
	9	18.8	8.6	50	9	AL628983
	10	18.8	8.6	50	9	AUI02832
	11	18.6	8.5	50	9	AUI06276
C	12	18.6	8.5	50	9	AUI06279
C	13	18.4	8.4	44	29	AL945796
C	14	18.2	8.3	45	10	BE882963
C	15	18.2	8.3	45	29	CG718854
C	16	18.2	8.3	48	28	AZ486468
C	17	18.2	8.3	50	9	AUI04187
C	18	18	8.2	41	28	CC326813
C	19	18	8.2	43	28	BH906713
C	20	17.8	8.1	50	9	AUI04910
C	21	17.6	8.0	40	9	AI634441
C	22	17.6	8.0	46	29	CC887026
C	23	17.6	8.0	49	12	BQ065334
C	24	17.4	7.9	28	13	BQ593079
C	25	17.4	7.9	45	12	BI218941
C	26	17.4	7.9	50	9	AUI05889
C	27	17.2	7.9	33	28	AZ840354
C	28	17.2	7.9	38	28	BZ292435
C	29	17.2	7.9	38	29	TA117B04P
C	30	17.2	7.9	41	28	AZ789628
C	31	17.2	7.9	50	9	AUI04911
C	32	17.2	7.9	50	9	AUI06272
C	33	17.2	7.9	50	9	AUI06275
C	34	17	7.8	43	28	BZ596218
C	35	17	7.8	49	28	BZ380372
C	36	17	7.8	50	9	AUI02511
C	37	17	7.8	50	9	AUI02514
C	38	17	7.8	50	9	AUI02516
C	39	17	7.8	50	9	AUI02517
C	40	17	7.8	50	9	AUI06287
C	41	16.8	7.7	37	28	BH418190
C	42	16.8	7.7	49	28	BZ664689
C	43	16.8	7.7	49	28	BZ664691
C	44	16.8	7.7	50	9	AUI06277
C	45	16.6	7.6	32	28	AZ332090

#### ALIGNMENTS

RESULT 1  
AZ481318  
LOCUS 50 bp DNA linear GSS 04-OCT-2000  
DEFINITION 1M0303B05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0303B05 R, genomic survey sequence.  
ACCESSION AZ481318  
VERSION AZ481318.1 GI:10642383  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 50)  
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weiss, R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Em' 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177



vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGG and 3' end primer CGACTCGAGTCGAGACA."

ORIGIN  
Query Match 9.5%; Score 20.8; DB 9; Length 43;  
Best Local Similarity 78.1%; Pred. No. 6e+04; 7; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 46 TCGCAGCTTCTCCCTCCAGCAATAAAGTGT 77  
Db 35 TGGTCCCTTCTTCTCCCAACAAAAATGGGT 4

RESULT 4  
LOCUS AU107911/c 50 bp mRNA linear EST 30-AUG-2001  
DEFINITION AU107911 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone ZR61770, mRNA sequence.

ACCESSION AU107911  
VERSION AU107911.1 GI:13557433  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites

JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929

COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES  
source  
1. .50  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ZR61770"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN  
Query Match 9.5%; Score 20.8; DB 9; Length 50;  
Best Local Similarity 64.6%; Pred. No. 6.5e+04;  
Matches 31; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 4 CTCCTGAGGAGCGGAGTGACCATCGGCGGAGAGCGAGTTCGAC 51  
Db 50 CACAGGACGGACGGGACCGGACCGCGGCTGCGCGGGGTTGGAC 3

RESULT 5  
LOCUS AU105891/c 50 bp mRNA linear EST 30-AUG-2001  
DEFINITION AU105891 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HTLB0018, mRNA sequence.

ACCESSION AU105891  
VERSION AU105891.1 GI:13555412

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE

JOURNAL  
MEDLINE  
PUBMED

COMMENT

Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES  
source

1. .50  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HTLB0018"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 9.2%; Score 20.2; DB 9; Length 50;  
Best Local Similarity 65.1%; Pred. No. 9.9e+04;  
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 6 CTGAGGAGCGGAGTGACCATCGGCGGAGAGAGGTTGC 48  
Db 48 CCAGAGCGGAGCGGAGAGCCCGGGGAGAGAGGGGNGC 6

RESULT 6  
LOCUS AU102843

DEFINITION AU102843 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone CAS11159, mRNA sequence.

ACCESSION AU102843  
VERSION AU102843.1 GI:13552364

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.

TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites

JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929

COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES  
source

1. .50  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HTLB0018"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 8.9%; Score 19.4; DB 9; Length 50;  
Best Local Similarity 70.3%; Pred. No. 1.7e+05;  
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 106 GATGAAAAATCAGGTGACACTGGAAGATACCA 142  
DB 5 GCTATAAAACAGCGAGGAGAACTGGCAGATACCA 41

RESULT 7  
LOCUS AUI02844 50 bp mRNA linear EST 30-AUG-2001  
DEFINITION AUI02844 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
CAS11168, mRNA sequence.  
ACCESSION AUI02844  
VERSION AUI02844.1 GI:13552365  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
Hata, H., Ota, T., Isegai, T., Tanaka, T., Morishita, S., Okubo, K.,  
Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
JOURNAL ENBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitono-Nakagawa, K., Maruyama, K., Suyama, A. and  
Sugano, S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),  
149-156 (1997).

FEATURES  
source  
1..50  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CAS11168"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 8.9%; Score 19.4; DB 9; Length 50;  
Best Local Similarity 70.3%; Pred. No. 1.7e+05;  
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 106 GATGAAAAATCAGGTGACACTGGAAGATACCA 142  
DB 5 GCTATAAAACAGCGAGGAGAACTGGCAGATACCA 41

RESULT 8  
LOCUS AAI37304 40 bp mRNA linear EST 09-FEB-1997  
DEFINITION mq80b07.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA  
clone IMAGE:585013 5' similar to TR:G1002369 G1002369 COATOMER  
PROTEIN. i, mRNA sequence.  
ACCESSION AAI37304  
VERSION AAI37304.1 GI:1700284  
KEYWORDS EST.

SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 40)  
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Marra M/Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:359661  
Possible reversed clone: similarity on wrong strand  
Seq primer: -28ml3 rev1 ET from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers  
1..40  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:585013"  
/tissue\_type="melanoma"  
/dev\_stage="M2 cells"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_lib="Stratagene mouse melanoma (#937312)"  
/note="Organ: skin; Vector: pBluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. From M2 cells, a highly metastatic derivative of  
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;  
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG  
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

ORIGIN

Query Match 8.6%; Score 18.8; DB 9; Length 40;  
Best Local Similarity 76.7%; Pred. No. 2.4e+05;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 171 GAAGTTTCTTAAGAGCAGACATGCCCTTT 200  
DB 5 GCAGGTTCTTTTGAACAGCCATGCGCTTT 34

RESULT 9  
LOCUS AL628983 50 bp mRNA linear EST 19-NOV-2003  
DEFINITION AL628983 XGC-gastrula Silurana tropicalis cDNA clone TGas011g16 5',  
mRNA sequence.  
ACCESSION AL628983  
VERSION AL628983.1 GI:16598466  
KEYWORDS EST.  
SOURCE Silurana tropicalis (western clawed frog)  
ORGANISM Silurana tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
Xenopodinae; Silurana.  
REFERENCE 1 (bases 1 to 50)  
AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.  
TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
JOURNAL Unpublished (2003)  
COMMENT Contact: Huckle E  
Sanger Institute  
Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: trop@sanger.ac.uk

This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.  
cDNA was oligo dt primed from 5ug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end  
Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
Host: Escherichia coli XL1-blue  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS SEQUENCE ID: TGAS011g16.plksp6  
Sequencing primer: Sp6.

FEATURES  
source

Location/Qualifiers  
1..50

/organism="Silurana tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="TGAS011g16"  
/dev\_stage="gastrula (stages 10.5-12 mixed)"  
/lab\_host="Escherichia coli XL1-blue"  
/clone\_lib="XGC-gastrula"  
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 8.6%; Score 18.8; DB 9; Length 50;  
Best Local Similarity 68.4%; Pred. No. 2.7e+05;  
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 CTCCTGAGGAGCGGAGTGCACCTCGGGCGGAGACG 41  
Db 13 CTCCTGTTCAAGGGGAGAGACTCGGGCGGAGCATG 50

RESULT 10  
AUI02832

LOCUS AUI02832 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
DEFINITION CAS01996, mRNA sequence.

ACCESSION AUI02832  
VERSION AUI02832.1 GI:13552353  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites

JOURNAL MEDLINE

PUBMED 21270072  
COMMENT EMBO Rep. 2 (5), 388-393 (2001)

CONTACT

Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES  
source

Location/Qualifiers  
1..50

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CAS01996"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 8.6%; Score 18.8; DB 9; Length 50;  
Best Local Similarity 76.7%; Pred. No. 2.7e+05;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 113 ATCAGTCAAGTGCACCTGGAAGATACCA 142  
Db 10 AAACAGCGAGGAGACACTGGCAGATACCA 39

RESULT 11  
AUI06276/c

LOCUS AUI06276 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
DEFINITION HRC04067, mRNA sequence.

ACCESSION AUI06276  
VERSION AUI06276.1 GI:13555797  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites

JOURNAL MEDLINE

PUBMED 21270072  
COMMENT EMBO Rep. 2 (5), 388-393 (2001)

CONTACT

Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES  
source

Location/Qualifiers  
1..50

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRC04067"  
/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 8.5%; Score 18.6; DB 9; Length 50;  
Best Local Similarity 65.9%; Pred. No. 3.1e+05;  
Matches 27; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 8 TGAGGAGCGGAGTGCACCTCGGGCGGAGAGGTTGC 48  
Db 48 TCGGAGACAGCGCGGAGGAGCGCGCGCGCGCGCGTGC 8

RESULT 12  
AUI06279/c

LOCUS AUI06279 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
DEFINITION HRC06165, mRNA sequence.

ACCESSION AUI06279  
VERSION AUI06279.1 GI:13555800  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,

```

/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-291H10-015349"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

```

```

ORIGIN
Query Match      8.4%; Score 18.4; DB 29; Length 44;
Best Local Similarity 69.4%; Pred. NO. 3.3e+05;
Matches 25; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      156 AGTGATTAAACAGCTGAGGTTGTTAAGAACGACGAC 191
      ||||| ||||| ||||| ||||| ||||| |||||
DB      43  AGCGATAACATAGGAAGGCTAATAGGAAGAAGGC 8

RESULT 14
BE882563
LOCUS
DEFINITION
601505866F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3907528 5',
mRNA sequence.
ACCESSION BE882963
BE882963
linear EST 20-OCT-2000

```

EST.	
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1. (bases 1 to 44)
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM9718 row: c column: 17 High quality sequence stop: 44. Location/Qualifiers
FEATURES	

```

source
1. 44
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:3907528"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 71"
/name="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN

Query Match      8.3%   Score 18.2;   DB 10;   Length 44;
Best Local Similarity 66.7%;   Pred. No. 3.8e+05;
Matches 26;   Conservative 0;   Mismatches 13;   Indels 0;   Gaps 0;

          7  CTGAGGAAGCGGGGAGTGGACCATCGGGCGGAGACGAGGT 45
          4  CTGAGGCGAGAGGTGGGACCATCTGGGGCCCTCGCGCTCT 42

```

RESULT 15  
CG718854/c

LOCUS 45 bp DNA linear GSS 20-OCT-2003

DEFINITION 1119054G10.1EL\_y1 1119 - RescueMu Grid AA Zea mays genomic, genomic survey sequence.

ACCESSION CG718854

VERSION CG718854.1 GI:37750174

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 45)

TITLE Walbot,V.

JOURNAL Maize genomic sequences found using engineered RescueMu transposon Unpublished (2001)

COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu

FEATURES source  
1..45  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="mixed background W23/A188/B73/K55"  
/db\_xref="taxon:4577"  
/tissue\_type="leaf"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 8.3%; Score 18.2; DB 29; Length 45;  
Best Local Similarity 66.7%; Pred. No. 3.8e+05;  
Matches 26; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 44 GTTGGACCTTTCCTCCCGCAATAAATCGTCTCG 82

DB 45 GATCCTCCCTCTCCATCAAGAGATATGAATGGTCTCG 7

Search completed: May 26, 2004, 09:17:45  
Job time : 1524.36 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:10:28 / Search time 982.615 Seconds  
(without alignments)  
5822.506 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_997\_1128

Perfect score: 132

Sequence: 1 acatgcacatcgcagga.....ctaccgcgcgtgcccg 132

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1603530

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hcg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rtd.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20.2	15.3	47	6	AR290994	AR290994 Sequence
2	19.6	14.8	48	6	AX611835	AX611835 Sequence
3	19.2	14.5	47	6	I07024	I07024 Sequence 6
4	19.2	14.5	47	6	I27626	I27626 Sequence 18
5	19.1	14.4	47	6	I07021	I07021 Sequence 3
6	18.2	13.8	47	6	I27623	I27623 Sequence 15
7	18.2	13.8	36	6	AX740373	AX740373 Sequence
8	18.2	13.8	41	6	A01533	A01533 Joining reg
9	18.2	13.8	48	6	AR051077	AR051077 Sequence
10	18.2	13.8	48	6	AR430291	AR430291 Sequence
11	18.1	13.6	45	6	AR078468	AR078468 Sequence
12	18.1	13.6	47	6	A26491	A26491 Synthetic I
13	18.1	13.6	47	6	AR118511	AR118511 Sequence
14	18.1	13.6	47	6	AR153892	AR153892 Sequence
15	18.1	13.6	48	6	AX611836	AX611836 Sequence
16	17.6	13.3	42	6	E16421	E16421 DNA encodin
17	17.6	13.3	45	6	E16414	E16414 DNA encodin
18	17.6	13.3	38	6	E16420	E16420 DNA encodin
19	17.4	13.2	32	6	AR373963	AR373963 Sequence
20	17.4	13.2	40	6	AR363937	AR363937 Sequence
21	17.2	13.0	43	4	OCT007	V00898 Rabbit trop
22	17.2	13.0	45	6	E16424	E16424 DNA encodin
23	17.2	13.0	45	6	E16427	E16427 DNA encodin
24	17.2	13.0	46	6	AR363127	AR363127 Sequence
25	17.2	13.0	46	6	AX769580	AX769580 Sequence
26	17.1	12.9	36	6	I91771	I91771 Sequence 5
27	17.1	12.9	47	6	AR289000	AR289000 Sequence
28	17.1	12.9	47	6	AX378744	AX378744 Sequence
29	17.1	12.9	48	6	AR031652	AR031652 Sequence
30	17.1	12.9	48	6	AR032616	AR032616 Sequence
31	17.1	12.9	48	6	I23356	I23356 Sequence 22
32	17.1	12.9	48	6	I90270	I90270 Sequence 11
33	17.1	12.9	48	6	I91030	I91030 Sequence 22
34	17.1	12.9	48	6	AR209280	AR209280 Sequence
35	16.8	12.7	30	6	I89979	I89979 Sequence 27
36	16.8	12.7	39	6	A22325	A22325 Primer O-2
37	16.8	12.7	39	6	A22326	A22326 Primer O-2
38	16.8	12.7	40	6	AR253352	AR253352 Sequence
39	16.8	12.7	40	6	BD138012	BD138012 Lectomedi
40	16.8	12.7	47	6	AR001266	AR001266 Sequence
41	16.8	12.7	47	6	I23424	I23424 Sequence 17
42	16.8	12.7	48	6	BD182587	BD182587 Anti TRAI
43	16.8	12.7	50	6	AR032939	AR032939 Sequence
44	16.8	12.7	50	6	E50843	E50843 Process for
45	16.8	12.7	50	6	I29679	I29679 Sequence 55

# ALIGNMENTS

RESULT 1	AR290994	AR290994	47 bp	DNA	linear	PAT 12-JUN-2003
LOCUS	Sequence	2729 from patent US 6537751.				
DEFINITION	AR290994					
ACCESSION	AR290994.1	GI:31678278				
VERSION						
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 47)					
AUTHORS	Cohen, D., Chumakov, I. and Blumenfeld, M.					
TITLE	Biallelic markers for use in constructing a high density					
JOURNAL	Disequilibrium map of the human genome					
	Patent: US 6537751-A 2729 25-MAR-2003;					

74 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCC 113

```
REFERENCE
AUTHORS      1 (bases 1 to 47)
TITLE        Capon,D.J. and Gregory,T.J.
JOURNAL      Adhesion variants
FEATURES     Patent: US 5565335-A 15 15-OCT-1996;
              Location/Qualifiers
              source
              1..47
              /organism="unknown"
              /mol_type="unassigned DNA"
ORIGIN
Query Match      14.4%; Score 19; DB 6; Length 47;
Best Local Similarity 71.4%; Pred. No. 2.6e+06;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 74 CTTGCTACTCGGCTGGATGAGCGCTGCTCCCTG 108
Db 12 CTTGCTTGTGTGCTGCAACTGGCGCTCCTCCAG 46
RESULT 7
AX740373/c
LOCUS          AX740373          36 bp      DNA      linear      PAT 10-MAY-2003
DEFINITION     Sequence 107 from Patent EP1300419.
ACCESSION      AX740373
VERSION        AX740373.1 GI:30523546
KEYWORDS       synthetic construct
SOURCE         synthetic construct
ORGANISM       artificial sequences.
REFERENCE
AUTHORS      1 Buettner,C., Schwarz,M., Knackmuss,S., Peter,K., Roettgen,P. and
              Little,M.
TITLE        Antibody of human origin for inhibiting thrombocyte aggregation
JOURNAL      Patent: EP 1300419-A 107 09-APR-2003;
              Affimed Therapeutics AG (DE)
FEATURES     Location/Qualifiers
              source
              1..36
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
              /note="Oligonucleotide"
ORIGIN
Query Match      13.8%; Score 18.2; DB 6; Length 36;
Best Local Similarity 87.0%; Pred. No. 4.3e+06;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 46 CAGCCCTGCATGCACACGTTCTG 68
Db 27 CAGCCCTGCTGTACGCGTTCTG 5
RESULT 8
A01533
LOCUS          A01533          41 bp      DNA      linear      PAT 25-FEB-1993
DEFINITION     Joining region of the bacterial origin and part of pBR322 origin in
              pBB9b28.
ACCESSION      A01533
VERSION        A01533.1 GI:344373
KEYWORDS       synthetic construct
SOURCE         synthetic construct
ORGANISM       artificial sequences.
REFERENCE
AUTHORS      1 (bases 1 to 41)
TITLE        EXPRESSION VECTORS AND METHOD FOR THEIR CONSTRUCTION
JOURNAL      Patent: WO 8902466-A 4 23-MAR-1989;
              Location/Qualifiers
              source
              1..41
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
ORIGIN
Query Match      13.8%; Score 18.2; DB 6; Length 48;
Best Local Similarity 74.2%; Pred. No. 4.2e+06;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 40 AGTTTGCAGCCCTGCTGATGCACACGTTCTGCG 70
Db 10 AGTATGCAACCCCTGCTGCGCGCTTTCGTTG 40
RESULT 9
AR051077
LOCUS          AR051077          48 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION     Sequence 20 from patent US 5830650.
ACCESSION      AR051077
VERSION        AR051077.1 GI:5974441
KEYWORDS       Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE
AUTHORS      1 (bases 1 to 48)
TITLE        Crea,R.
JOURNAL      Walk-through mutagenesis
FEATURES     Patent: US 5830650-A 20 03-NOV-1998;
              Location/Qualifiers
              source
              1..48
              /organism="unknown"
              /mol_type="unassigned DNA"
ORIGIN
Query Match      13.8%; Score 18.2; DB 6; Length 48;
Best Local Similarity 48.9%; Pred. No. 4e+06;
Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
QY 64 TTCTGCGCGGCTTGCTACTCGGCTGGATGGAGCGCTCCTCCTGTG 110
Db 1 TACTGCGCGGTARCTMTCTMTGTCAGCASTTGTCTCTCKMCKYTTG 47
RESULT 10
AR430291
LOCUS          AR430291          48 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION     Sequence 20 from patent US 6649340.
ACCESSION      AR430291
VERSION        AR430291.1 GI:40191060
KEYWORDS       Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE
AUTHORS      1 (bases 1 to 48)
TITLE        Crea,R.
JOURNAL      Walk-through mutagenesis
FEATURES     Patent: US 6649340-A 20 18-NOV-2003;
              Location/Qualifiers
              source
              1..48
              /organism="unknown"
              /mol_type="genomic DNA"
ORIGIN
Query Match      13.8%; Score 18.2; DB 6; Length 48;
Best Local Similarity 48.9%; Pred. No. 4e+06;
Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
QY 64 TTCTGCGCGGCTTGCTACTCGGCTGGATGGAGCGCTCCTCCTGTG 110
Db 1 TACTGCGCGGTARCTMTCTMTGTCAGCASTTGTCTCTCKMCKYTTG 47
RESULT 11
AR078468/c
LOCUS          AR078468          45 bp      DNA      linear      PAT 31-AUG-2000
DEFINITION     Sequence 8 from patent US 5962664.
ACCESSION      AR078468
```

```

VERSION      AR078468.1  GI:10005214
KEYWORDS
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 45)
AUTHORS      Friedhoff, A.J., Baeham, D.A. and Miller, J.C.
TITLE        Psychosis protecting nucleic acid, peptides, compositions and
              method of use
JOURNAL      Patent: US 5962664-A 8 05-OCT-1999;
FEATURES     Location/Qualifiers
              source
              1..45
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN
Query Match      13.6%; Score 18; DB 6; Length 45;
Best Local Similarity 80.8%; Pred. No. 4.6e+06;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 74 CTTGCTACTCGGCTTGCTACTCGGGCTGGATGGAGCGC 99
      |||||  |||||  |||||  |||||  |||||  |||||
Db 29 CTTGTGCCAGGCTGGATGGAGTGC 4

RESULT 12
LOCUS      A26491
DEFINITION Synthetic IDUA probe ID47.
ACCESSION  A26491
VERSION     A26491.1  GI:11248227
KEYWORDS   synthetic construct
           synthetic construct
           artificial sequences.
ORGANISM   1 (bases 1 to 47)
REFERENCE  1 (bases 1 to 47)
AUTHORS    Hopwood, J.J., Orsborn, A.M., Anson, D.S., Clements, P.R., Morris, C.P.,
           Nelson, P.V. and Scott, H.S.
TITLE      Synthetic alpha-L-iduronidase and genetic sequences encoding same
JOURNAL    Patent: WO 9310244-A 1 27-MAY-1993;
           WOMENS & CHILDRENS HOSPITAL (AU)
FEATURES   Location/Qualifiers
           source
           1..47
           /organism="synthetic construct"
           /mol_type="unassigned DNA"
           /db_xref="taxon:32630"

ORIGIN
Query Match      13.6%; Score 18; DB 6; Length 47;
Best Local Similarity 64.3%; Pred. No. 4.5e+06;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 61 ACGTTCTGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
      |||||  |||||  |||||  |||||  |||||  |||||
Db 47 ACGTTGTCGAAGTCGTGCTGCTGGCTCGGCTCGTTCCAGGCTCG 6

RESULT 13
LOCUS      AR118511/c
DEFINITION Sequence 1 from patent US 6149909.
ACCESSION  AR118511
VERSION     AR118511.1  GI:14100421
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 47)
AUTHORS    Scott, H.Steele., Anson, D.Stewart., Orsborn, A.Marie.,
           Nelson, P.Victor., Clements, P.Roy., Morris, C.Phillip. and
           Hopwood, J.Joseph.
TITLE      Synthetic .alpha.-L-iduronidase and genetic sequences encoding same
JOURNAL    Patent: US 6149909-A 1 21-NOV-2000;

```

```

FEATURES     Location/Qualifiers
           source
           1..47
           /organism="unknown"
           /mol_type="unassigned DNA"

ORIGIN
Query Match      13.6%; Score 18; DB 6; Length 47;
Best Local Similarity 64.3%; Pred. No. 4.5e+06;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 61 ACGTTCTGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
      |||||  |||||  |||||  |||||  |||||  |||||
Db 47 ACGTTGTCGAAGTCGTGCTGCTGGCTCGGCTCGTTCCAGGCTCG 6

RESULT 14
LOCUS      AR153892/c
DEFINITION Sequence 1 from patent US 6238662.
ACCESSION  AR153892
VERSION     AR153892.1  GI:15121945
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 47)
AUTHORS    Scott, H.Steel., Anson, D.Stewart., Orsborn, A.Marie.,
           Nelson, P.Victor., Clements, P.Roy., Morris, C.Phillip. and
           Hopwood, J.Joseph.
TITLE      Synthetic .alpha.-L-iduronidase and genetic sequences encoding same
JOURNAL    Patent: US 6238662-A 1 29-MAY-2001;
FEATURES   Location/Qualifiers
           source
           1..47
           /organism="unknown"
           /mol_type="unassigned DNA"

ORIGIN
Query Match      13.6%; Score 18; DB 6; Length 47;
Best Local Similarity 64.3%; Pred. No. 4.5e+06;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 61 ACGTTCTGCGGCTTGCTACTCGGGCTGGATGGAGCGCTCG 102
      |||||  |||||  |||||  |||||  |||||  |||||
Db 47 ACGTTGTCGAAGTCGTGCTGCTGGCTCGGCTCGTTCCAGGCTCG 6

RESULT 15
LOCUS      AX611836
DEFINITION Sequence 2861 from Patent WO02072882.
ACCESSION  AX611836
VERSION     AX611836.1  GI:28407265
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS    Cullen, P. and Seedorf, U.
TITLE      Coronary chip
JOURNAL    Patent: WO 02072882-A 2861 19-SEP-2002;
           OGHAM GmbH (DE)
FEATURES   Location/Qualifiers
           source
           1..48
           /organism="Homo sapiens"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9606"

ORIGIN
Query Match      13.6%; Score 18; DB 6; Length 48;
Best Local Similarity 70.6%; Pred. No. 4.5e+06;
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Qy 23 TGCTGCACGACTGGCTGAGTTTGACAGCCCTGCAT 56  
Db 12 TGATAAATGACTTCAAGAGCTTGGAGCCCGGCAT 45

Search completed: May 26, 2004, 07:39:18  
Job time : 984.615 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:01:38 ; Search time 111.477 Seconds  
(without alignments)  
5030.293 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_997\_1128

Perfect score: 132  
Sequence: 1 acatgcacatcgtccagga.....ctaccgcccgtgcgcgtg 132

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3185356

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	20.2	15.3	42	6	ABL95873	ABL95873 LLPL-rela
2	19.2	14.5	50	9	ADD31850	Add31850 Synthetic
3	19.2	14.4	47	3	Az268382	Az268382 Human map
4	18.8	14.2	41	6	ABN83043	ABN83043 Human ank
5	18.2	13.8	36	9	AD803377	AD803377 Human inn
6	18.2	13.8	48	2	AAV60897	AAV60897 Oligonuci
7	18.2	13.8	48	9	ADD54968	Add54968 Heavy cha
8	18	13.6	47	2	AAQ42415	AAQ42415 IDUA prim
9	18	13.6	48	7	AB56353	AB56353 PCR prime
10	18	13.6	50	6	AB201079	AB201079 Human leu
11	17.8	13.5	50	4	AAU31778	AAU31778 Human SNP
12	17.6	13.3	42	2	AAV02057	AAV02057 Mutated h
13	17.6	13.3	45	2	AAV02046	AAV02046 hCS-15 pe
14	17.6	13.3	48	2	AAV02056	AAV02056 Mutated h
15	17.6	13.3	50	4	AAU34532	AAU34532 Human SNP
16	17.4	13.2	28	5	AAFC0356	AAFC0356 Human che
17	17.4	13.2	49	8	ACF04078	ACF04078 Cdc20 sup
18	17.2	13.0	41	6	ABN83042	ABN83042 Human ank
19	17.2	13.0	41	6	ABN75713	ABN75713 Tumour su
20	17.2	13.0	45	2	AAV02060	AAV02060 Mutated h
21	17.2	13.0	45	2	AAV02063	AAV02063 Mutated h
22	17.2	13.0	46	1	AAN60796	AAN60796 Sequence
23	17.2	13.0	46	2	AAT28216	AAT28216 Feline he

C	24	17.2	13.0	46	7	ABZ68916	ABZ68916 PCR prime
	25	17.2	13.0	50	4	AAU33784	AAU33784 Human SNP
	26	17	12.9	36	2	AAQ97839	AAQ97839 Human EI
	27	17	12.9	36	2	AAU41762	AAU41762 Ubiquitin
	28	17	12.9	36	2	AAU78503	AAU78503 Human EI
	29	17	12.9	36	3	AAU61629	AAU61629 Human EI
C	30	17	12.9	47	3	AAZ66388	AAZ66388 Human map
C	31	17	12.9	47	6	ABK41285	ABK41285 Human LSR
	32	17	12.9	48	2	AAQ69478	AAQ69478 Human leu
	33	17	12.9	48	2	AAU63940	AAU63940 Human LFA
	34	17	12.9	48	2	AAU17228	AAU17228 Test sequ
	35	17	12.9	48	6	ABK82719	ABK82719 DNA bindi
	36	17	12.9	48	10	ADE80258	ADE80258 Duplex ol
C	37	17	12.9	50	2	AAV07319	AAV07319 Nucleotid
C	38	17	12.9	50	4	AAU33674	AAU33674 Human SNP
C	39	17	12.9	50	4	AAU30334	AAU30334 Human SNP
C	40	16.8	12.7	38	4	AAU54939	AAU54939 PCR prime
	41	16.8	12.7	40	2	AAZ28006	AAZ28006 Human lec
	42	16.8	12.7	40	3	AAU96033	AAU96033 Human lec
C	43	16.8	12.7	47	2	AAQ28991	AAQ28991 Sequence
	44	16.8	12.7	47	3	AAZ67208	AAZ67208 Human map
	45	16.8	12.7	47	3	AAZ68974	AAZ68974 Human map

ALIGNMENTS

RESULT 1

ABL95873  
ID ABL95873 standard; DNA; 42 BP.

XX ABL95873;

DT 19-JUN-2002 (first entry)

DE LLPL-related oligonucleotide #2.

KW Recombinant protein production; drug; reagent; food stuff; ss.

OS Unidentified.

XX WO200208417-A1.

XX 31-JAN-2002.

XX 25-JUL-2001; 2001WO-JP006392.

XX 25-JUL-2000; 2000JP-00229064.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Ito T, Tanaka Y, Kondo M;

XX WPI; 2002-179906/23.

PT Production of recombinant proteins in prokaryotes or eukaryotes particularly with target proteins obtainable through gene recombination technique, for use as drugs, reagents, raw materials for industries and feeding stuffs.

XX Example 23; Fig 29; 137pp; Japanese.

PS The present invention relates to a method for producing recombinant proteins. The method comprises preparing a recombinant vector for transforming a host cell before culturing the obtained transformant, assaying expression of the reporter gene and confirming high expression of the reporter gene. The recombinant proteins are useful as drugs, reagents, raw materials for industries and feeding stuffs. Also, the proteins are obtainable on large-scale production. The present sequence was used to illustrate the invention

XX Sequence 42 BP; 3 A; 16 C; 16 G; 7 T; 0 U; 0 Other;

Query Match 15.3%; Score 20.2; DB 6; Length 42;  
Best Local Similarity 68.3%; Pred. No. 1.6e+04;  
Matches 28; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 10 ATCTGCCAGGACCTGCTGCACGACTGCGTGAGTTTGACGCC 50  
1 ATGGGCTGCACCTGCGGCTTACCGGCTGGGCTGCTGCC 41

Db

RESULT 2  
ADD31850/C  
ID ADD31850 standard; DNA; 50 BP.  
AC ADD31850;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Synthetic antibody oligonucleotide SEQ ID NO:89.  
XX recombination product; synthetic gene technology; antibody; ss.  
XX Synthetic.  
XX WO2003064611-A2.  
XX  
XX 07-AUG-2003.  
XX  
XX 29-JAN-2003; 2003WO-US002612.  
XX  
XX 30-JAN-2002; 2002US-00062188.  
XX (EGEA-) EGEA BIOSCIENCES INC.  
XX  
XX Evans GA;  
XX  
XX WPI; 2003-663477/62.  
XX  
XX Creating recombination products between two distinct nucleotide  
XX sequences, useful in the field of synthetic gene technology, and in  
XX assembling a library, or a population or a collection of polypeptide  
XX variants.

XX Example 2; SEQ ID NO 89; 132pp; English.

XX The present invention describes a method for creating a collection of  
XX recombination products between two nucleotide sequences. The method  
XX comprises combining an initial set of oligonucleotides corresponding to a  
XX first nucleotide sequence with a subsequent set of oligonucleotides  
XX corresponding to a distinct nucleotide sequence and further combining the  
XX initial and subsequent sets of combination oligonucleotides having a  
XX sequence region corresponding to the initial nucleotide sequence and a  
XX sequence region corresponding to the second oligonucleotide sequence.  
XX Also described is a method of creating a collection of recombination  
XX products between two genes. The methods and compositions of the present  
XX invention are useful in the field of synthetic gene technology, and more  
XX specifically, to generating a collection of recombination products  
XX between distinct nucleotide sequences. They can also be used in  
XX assembling a library, or a population or a collection of polypeptide  
XX variants that correspond to single or multiple polynucleotide  
XX recombination products. The present sequence represents an  
XX oligonucleotide used in the assembly of a synthetic antibody, which is  
XX used in an example from the present invention.

XX Sequence 50 BP; 12 A; 11 C; 18 G; 9 T; 0 U; 0 Other;

Query Match 14.5%; Score 19.2; DB 9; Length 50;  
Best Local Similarity 75.0%; Pred. No. 3.2e+04;  
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 50 CCTGCATGCACAGCTTCTGCGGCTTGCTTAC 81  
44 CCTGACGTACCACTGTTCCGCGCGCTTCTTAC 13

Db

RESULT 3  
AAZ68382  
ID AAZ68382 standard; DNA; 47 BP.  
XX  
XX AAZ68382;  
XX  
XX 10-SEP-2001 (first entry)  
XX  
XX Human map-related biallelic marker SEQ ID NO:2729.  
XX  
XX Human genome; biallelic marker; high density disequilibrium map;  
XX genomic map; haplotype; phenotype; polymorphic base; genotyping;  
XX haplotyping; hybridisation; identification; characterisation; diagnosis;  
XX single nucleotide polymorphism; SNP; ds.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX variation replace(24,C)  
XX /\*tag= a  
XX /standard\_name= "single nucleotide polymorphism"  
XX  
XX WO9954500-A2.  
XX  
XX 28-OCT-1999.  
XX  
XX 21-APR-1999; 99WO-IB000822.  
XX  
XX 21-APR-1998; 98US-0082614P.  
XX 23-NOV-1998; 98US-0109732P.  
XX  
XX (GEST ) GENSET.  
XX  
XX Cohen D, Blumenfeld M, Chumakov I;  
XX  
XX WPI; 2000-013267/01.  
XX  
XX Novel biallelic markers used to construct a high density disequilibrium  
XX map of the human genome.  
XX  
XX Claim 3; Page 812; 2745pp; English.

XX AAZ65654 to AAZ69578 represent human biallelic markers from the present  
XX invention, which contain a polymorphic base at position 24 of their  
XX nucleotide sequences. AAZ69579 to AAZ77440 represent amplification  
XX primers for the biallelic markers. The biallelic markers of the invention  
XX have a variety of uses: they can be used for high density mapping of the  
XX human genome, and in complex association studies and haplotyping studies  
XX which are useful in determining the genetic basis for disease states.  
XX Compositions and methods of the invention can also be useful for the  
XX identification of the targets for the development of pharmaceutical  
XX agents and diagnostic methods, as well as the characterisation of the  
XX differential efficacious responses to and side effects from  
XX pharmaceutical agents acting on a disease as well as other treatment.  
XX N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and  
XX 3367, are not actually given a sequence in the Sequence Listing from the  
XX present invention

XX Sequence 47 BP; 11 A; 13 C; 10 G; 13 T; 0 U; 0 Other;

Query Match 14.4%; Score 19; DB 3; Length 47;  
Best Local Similarity 65.1%; Pred. No. 3.7e+04;  
Matches 28; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 17 AGGACCTGCTGCACGACTGCGTGAGTTTGACGCCCTGCATGCA 59  
5 AGGTCTCGACGACGATGTGTACCACTTTGAAAGCCTCCATCCA 47

Db

RESULT 4  
ABN83043  
ID ABN83043 standard; DNA; 41 BP.





antigen-binding region of the monoclonal antibody (MAb) MCP603 (which binds phosphocholine) is altered to contain the catalytic triad residues for a serine protease. Specifically the amino acids to be altered are selected from the Asp of the complementarity determining region (CDR) 1 region of the variable heavy chain (Vh) of the antibody, the His of Vh CDR3 and the Ser of the CDR2 from the light chain variable region (VL). The mutagenesis is by a "walk-through" method. Oligonucleotides MCV608A1-MCV609J are used to assemble the Vh gene for monoclonal antibody MCP603 for the construction of variants (see AAV60860-V60880)

Sequence 48 BP; 4 A; 11 C; 9 G; 12 T; 0 U; 12 Other;

Query Match 13.8%; Score 18.2; DB 2; Length 48;  
Best Local Similarity 48.9%; Pred. No. 6.3e+04;  
Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 64 TTCTGCGGGTTCCTACTCGGCTGTGATGAGCGTGCTGCCTGTG 110  
||||| :||| :||| :||| :||| :||| :||| :|||  
Db 1 TACTGCGGGCTARCTMCTMTRGACGASTTSGTMCTCKMKKYTTG 47

RESULT 7  
ADD54968  
IID ADD54968 standard; DNA; 48 BP.  
XX AC ADP54968;  
XX AC  
XX 15-JAN-2004 (first entry)  
DE Heavy chain variable region variant oligonucleotide #17.  
XX ss; mutagenesis; protein mutagenesis; screening.  
XX Synthetic.  
XX US2003194807-A1.  
XX 16-OCT-2003.  
XX 20-FEB-2003; 2003US-00371404.  
XX 02-NOV-1992; 92US-00930600.  
XX 30-MAY-1995; 95US-00453623.  
XX (CREA/) CREA R.  
PA Crea R;  
PI WPI; 2003-844460/78.  
XX Mutagenesis of a protein comprises introducing a predetermined amino acid  
XX into each set of selected sequence positions in a predefined region of  
XX the protein to produce a protein library comprising mutant proteins.

Example 1; Page 12; 44pp; English.

The invention relates to a method of mutagenesis of a protein. The methods are useful for generating libraries of mutant proteins that are of a practical size for screening, for studying the role of amino acids in protein structure and function and for developing new or improved proteins and polypeptides such as enzymes, antibodies their binding fragments or analogues. The present sequence is used in the exemplification of the invention.

Sequence 48 BP; 4 A; 11 C; 9 G; 12 T; 0 U; 12 Other;

Query Match 13.8%; Score 18.2; DB 9; Length 48;  
Best Local Similarity 48.9%; Pred. No. 6.3e+04;  
Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 64 TTCTGCGGGTTCCTACTCGGCTGTGATGAGCGTGCTGCCTGTG 110  
||||| :||| :||| :||| :||| :||| :||| :|||  
Db 1 TACTGCGGGCTARCTMCTMTRGACGASTTSGTMCTCKMKKYTTG 47

KW disseminated intravascular coagulation; stem cell; transplantation;  
KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant;  
KW consensus variable heavy chain domain; CONVH; PCR; ss; PCR knitting.  
XX Synthetic.  
XX WO200278612-A2.  
XX 10-OCT-2002.  
XX 02-APR-2002; 2002WO-US010301.  
XX 02-APR-2001; 2001US-0281183P.  
XX (PURD ) PURDUE PHARMA LP.  
XX Solitis DA, Burch RM, Ogert RA;  
XX WPI; 2003-040615/03.  
XX New thrombopoietin synthetoids, useful for stimulating proliferation,  
XX growth, or differentiation of hematopoietic cells, for treating or  
XX preventing hematopoietic or immune disorders, e.g. thrombocytopenia.  
XX Example 1; Page 72; 97pp; English.  
XX The invention discloses a variant of an immunoglobulin (Ig) variable  
XX heavy or light chain domain that comprises at least one complementarity  
XX determining region (CDR) and framework regions flanking the CDR. The CDR  
XX also has added or substituted to it, at least one binding sequence which  
XX is heterologous to the CDR and is an antigenic, agonistic sequence from a  
XX thrombopoietin (TPO) receptor (MPL) binding sequence. The antigenic  
XX sequence can be a binding sequence heterologous to the CDR, a cytotoxic T  
XX -lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper  
XX cell sequence or a combination of each. The variant or thrombopoietin  
XX synthetoid, pharmaceutical and vaccine compositions are useful for  
XX stimulating proliferation, growth or differentiation of hematopoietic  
XX cells, particularly platelet progenitor cells. The variants are also  
XX useful for treating or preventing hematopoietic or immune disorders  
XX resulting from chemotherapy, radiation therapy, or bone marrow  
XX transfusions (e.g. thrombocytopenia or disseminated intravascular  
XX coagulation). Compositions comprising the synthetoids can be used for  
XX the mobilisation, amplification and ex vivo expansion of stem cells and  
XX committed precursor cells for autologous and allogeneic transplantation  
XX as well as for the expansion of stem cells for gene therapy. They are  
XX also useful as diagnostic or analytical reagents for studying the  
XX function of thrombopoietin and its receptor in vivo or in vitro. The  
XX sequence presented is the PCR primer, TPOVHP6, which was used to  
XX construct a variable heavy chain region gene containing the MPL binding  
XX sequence of TPO using the PCR knitting technique  
XX  
SQ Sequence 48 BP; 10 A; 14 C; 15 G; 9 T; 0 U; 0 Other;  
Query Match 13.6%; Score 18; DB 7; Length 48;  
Best Local Similarity 64.3%; Pred. No. 7.2e+04;  
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 71 CGGCTTGCTACTCGGCTGATGAGCGCTCGTCCCTGTGTC 112  
DB 48 CTGTTTACTGCGCTAGGATCGAGGCGCCACCTTGAGAC 7  
RESULT 10  
ABZ01079  
ID ABZ01079 standard; DNA; 50 BP.  
XX ABZ01079;  
AC ABZ01079;  
XX 09-JAN-2003 (first entry)  
XX Human leukocyte gene expression profiling probe SEQ ID NO 1070.  
XX T7; leukocyte; gene expression profiling; allograft rejection;

KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;  
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;  
KW ss.  
XX Homo sapiens.  
XX WO200257414-A2.  
XX 25-JUL-2002.  
XX 22-OCT-2001; 2001WO-US047856.  
XX 20-OCT-2000; 2000US-0241994P.  
XX 08-JUN-2001; 2001US-0296764P.  
XX (BIOC-) BIOCARDIA INC.  
XX Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;  
XX Ly N, Woodward R, Quettermous T, Johnson F;  
XX WPI; 2002-636525/68.  
XX New system for leukocyte expression profiling, diagnosing a disease, or  
XX monitoring (the rate of) progression of a disease, e.g. atherosclerosis  
XX or congestive heart failure, comprises diagnostic oligonucleotides.  
XX Claim 1; Page 359; Opp; English.  
XX The invention relates to a system for detecting gene expression, which  
XX comprises one or two isolated DNA molecules that detect expression of a  
XX gene, where the gene corresponds to any of 8143 oligonucleotides  
XX (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful  
XX for leukocyte expression profiling. It is particularly useful for  
XX diagnosing a disease, monitoring (rate of) progression of a disease,  
XX predicting therapeutic outcome, determining prognosis for a patient,  
XX predicting disease complications in an individual or monitoring response  
XX to treatment in an individual. The diseases include cardiac allograft  
XX rejection, kidney allograft rejection, liver allograft rejection,  
XX atherosclerosis, congestive heart failure, systemic lupus erythematosus,  
XX rheumatoid arthritis, osteoarthritis or cytomegalovirus infection  
XX  
SQ Sequence 50 BP; 11 A; 12 C; 9 G; 18 T; 0 U; 0 Other;  
Query Match 13.6%; Score 18; DB 6; Length 50;  
Best Local Similarity 70.6%; Pred. No. 7.2e+04;  
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 51 CTGCATGCACAGTTCGTGCGGCTGTGCTACTCG 84  
DB 10 CTGTGTGAACCTGTTCTGCATGCTGTGCTAATGG 43  
RESULT 11  
AAL31778/c  
ID AAL31778 standard; DNA; 50 BP.  
XX AAL31778;  
XX 24-JAN-2002 (first entry)  
XX Human SNP oligonucleotide #4986.  
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.  
XX Homo sapiens.  
XX

```

PR 15-OCT-1996;    96JP-00272422.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Fukusumi S, Kitada C;
XX
XX WPI; 1998-042177/04.
DR DR P-PSDB; AAW44022.
XX
XX Peptide having corticostatin or somatostatin or activity - useful as
PT anticancer and antiulcer agent, and for control of dementia and growth
PT abnormalities.
XX
XX Claim 14; Page 138; 174pp; Japanese.
XX
XX This sequence encodes a mutated hCS peptide, which is a peptide of the
CC invention, and has corticostatin or somatostatin activity. Antibodies
CC recognising hCS-17 can be used to screen for a compound that modulates,
CC i.e. an agonist or antagonist, the binding of hCS-17 to its receptor, and
CC to assay for hCS-17, e.g. diagnosis. hCS-17, the DNA encoding it or a
CC receptor agonist or antagonist can be used to treat and prevent hormone-
CC producing tumours (e.g. tumours producing gastrin or insulin), gastric
CC ulcers and dementia, regulate sleep and control growth disorders (e.g.
CC acromegaly, gigantism and dwarfism). They can also be used to regulate
CC hormone secretion and the digestive system (e.g. to treat diabetes), and
CC as tumour multiplication or neural inhibitors.
XX
XX Sequence 42 BP; 7 A; 15 C; 8 G; 12 T; 0 U; 0 Other;
SQ
Query Match      13.3%; Score 17.6; DB 2; Length 42;
Best Local Similarity 65.0%; Pred.No. 9.1e+04;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 48 GCCTCGATGCACACTTCTGCGGGCTTGCTACTCGGC 87
Db 3 GCCTCGAGAAATTCTTCTGGAAGACCTTCCTCCTGCG 42
|||||
|||
|
RESULT 13
AAV02046
ID ID AAV02046 standard; cDNA; 45 BP.
XX
XX AAV02046;
XX
XX 26-JUN-1998 (first entry)
XX
XX hCS-15 peptide coding sequence.
XX
XX Corticostatin; somatostatin; hCS-15 peptide; hormone-producing tumour;
KW gastric ulcer; dementia; growth disorder; hormone secretion regulation;
KW digestive system regulation; neural inhibitor; therapy; ds.
XX
XX Synthetic.
XX
XX WO9746668-A1.
XX
XX 11-DEC-1997.
XX
XX 05-JUN-1997; 97WO-JP001911.
XX
XX 07-JUN-1996; 96JP-00146052.
PR PR 19-SEP-1996; 96JP-00247710.
PR PR 15-OCT-1996; 96JP-00272422.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
PA
XX Hinuma S, Fukusumi S, Kitada C;
XX
XX WPI; 1998-042177/04.
DR DR P-PSDB; AAW44019.
XX
XX Peptide having corticostatin or somatostatin or activity - useful as
PT anticancer and antiulcer agent, and for control of dementia and growth
PT abnormalities.

```



CC organisms

XX

SQ Sequence 50 BP; 12 A; 9 C; 14 G; 15 T; 0 U; 0 Other;

Query Match 13.3%; Score 17.6; DB 4; Length 50;

Best Local Similarity 65.0%; Pred. No. 9.4e+04;

Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 30 CGACTGCGTGAGTTTGCAGCCCTGCATGCGACACGTTCTTGC 69

Db 6 CAACTGTGTGAGTGTGATAGCTGCTGCCACCTGTTTGC 45

Search completed: May 26, 2004, 05:53:46  
Job time : 114.477 secs



CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236.311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444PIC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-236-311-18

Query Match 14.5%; Score 19.2; DB 1; Length 47;  
Best Local Similarity 67.5%; Pred. No. 4.5e+03;  
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 74 CTGTGCTACTCGGCTGGATGGAGCGCTGCTCCCTGTGTGCC 113  
Db 46 CTGTGCTTCTGCTGCTGCAACTGGCGCTCTCTCCAGCAGCC 7

RESULT 3  
US-08-457-918-18/c  
Sequence 18, Application US/08457918  
Patent No. 6117655  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457.918

FILING DATE: 1-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444PIC3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-457-918-18

Query Match 14.5%; Score 19.2; DB 3; Length 47;  
Best Local Similarity 67.5%; Pred. No. 4.5e+03;  
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 74 CTGTGCTACTCGGCTGGATGGAGCGCTGCTCCCTGTGTGCC 113  
Db 46 CTGTGCTTCTGCTGCTGCAACTGGCGCTCTCTCCAGCAGCC 7

RESULT 4  
US-08-236-311-15  
Sequence 15, Application US/08236311  
Patent No. 5565335  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236.311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/250785
/ FILING DATE: 28-SEP-1988
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/104329
/ FILING DATE: 02-OCT-1987
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hasak, Janet E.
/ REGISTRATION NUMBER: 28,616
/ REFERENCE/DOCKET NUMBER: 444PIC2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/225-1896
/ TELEFAX: 415/952-9881
/ TELEX: 910/371-7168
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 47 bases
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-236-311-15

Query Match 14.4%; Score 19; DB 1; Length 47;
Best Local Similarity 71.4%; Pred. No. 5.2e+03;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 74 CTTGCTACTCGGCTGGATGGAGCGCTGTCCTCG 108
Db 12 CTTGCTTCTGCTGCTGCACTGGCGCTCCCTCCAG 46

RESULT 5
US-08-457-918-15
/ Sequence 15, Application US/08457918
/ Patent No. 6117855
/ GENERAL INFORMATION:
/ APPLICANT: Capon, Daniel J.
/ APPLICANT: Gregory, Timothy J.
/ TITLE OF INVENTION: Adheson Variants
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 460 Point San Bruno Blvd
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: patin (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/457,918
/ FILING DATE: 1-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/236311
/ FILING DATE: 02-MAY-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/936190
/ FILING DATE: 26-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/842777
/ FILING DATE: 18-FEB-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/250785
/ FILING DATE: 28-SEP-1988
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/104329
/ FILING DATE: 02-OCT-1987
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kubinec, Jeffrey S.
```

```
/ REGISTRATION NUMBER: 36,575
/ REFERENCE/DOCKET NUMBER: P0444PIC3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/225-8228
/ TELEFAX: 415/952-9881
/ TELEX: 910/371-7168
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 47 bases
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-457-918-15

Query Match 14.4%; Score 19; DB 3; Length 47;
Best Local Similarity 71.4%; Pred. No. 5.2e+03;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 74 CTTGCTACTCGGCTGGATGGAGCGCTGTCCTCG 108
Db 12 CTTGCTTCTGCTGCTGCACTGGCGCTCCCTCCAG 46

RESULT 6
US-08-452-724A-20
/ Sequence 20, Application US/08452724A
/ Patent No. 5830650
/ GENERAL INFORMATION:
/ APPLICANT: Crea, Roberto
/ TITLE OF INVENTION: Walk-Through Mutagenesis
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
/ STREET: 2 Militia Drive
/ CITY: Lexington
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02173
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/452,724A
/ FILING DATE: 30-MAY-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/930,600
/ FILING DATE: 05-APR-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US91/02362
/ FILING DATE: 05-APR-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/505,314
/ FILING DATE: 05-APR-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brook Esq., David E.
/ REGISTRATION NUMBER: 22,592
/ REFERENCE/DOCKET NUMBER: RC90-01AZ
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 861-6240
/ TELEFAX: (617) 861-9540
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 48 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
US-08-452-724A-20

Query Match 13.8%; Score 18.2; DB 2; Length 48;
Best Local Similarity 48.9%; Pred. No. 8.8e+03;
```



```

Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
QY 64 TTCTCGCGGCTTCTACTCGGCTGATGGAGCGTCTGTCCTGTG 110
   ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1 TACTGCGCGGTARCTCTCTWTRGCAGCASTTSGTMTCTYCKYCKYTTG 47

RESULT 7
US-08-453-623-20
; Sequence 20, Application US/08453623
; Patent No. 6649340
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: Walk-Through Mutagenesis
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,623
; FILING DATE: 30-May-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,600
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: PCT/US91/02362
; FILING DATE: 05-APR-1991
; APPLICATION NUMBER: US 07/505,314
; FILING DATE: 05-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: RC90-01AY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-08-453-623-20

Query Match 13.8%; Score 18.2; DB 4; Length 48;
Best Local Similarity 48.9%; Pred. No. 8.8e+03;
Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
QY 64 TTCTCGCGGCTTCTACTCGGCTGATGGAGCGTCTGTCCTGTG 110
   ||||| :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1 TACTGCGCGGTARCTCTCTWTRGCAGCASTTSGTMTCTYCKYCKYTTG 47

RESULT 8
US-08-602-716A-8/c
; Sequence 8, Application US/08602716A
; Patent No. 5962664
; GENERAL INFORMATION:
; APPLICANT: FRIEDHOFF, Arnold J.
; APPLICANT: BASHAM, Daryl A.
; APPLICANT: MILLER, Jeanette C.
; TITLE OF INVENTION: PSYCHOSIS PROTECTING NUCLEIC ACID,
; TITLE OF INVENTION: PEPTIDES, COMPOSITIONS AND METHODS OF USE

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/191,171  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/494,104  
FILING DATE: 23-JUN-1995  
APPLICATION NUMBER: AU PK9490/91  
FILING DATE: 14-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PCT/AU92/00611  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/084,254  
FILING DATE: 07-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8978Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Oligonucleotide DNA  
US-09-191-171-1

Query Match 13.6%; Score 18; DB 3; Length 47;  
Best Local Similarity 64.3%; Pred. No. 1e+04;  
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 61 ACCTTCTCGCGGCTGCTACTCGGCTGGATGAGCGCTCG 102  
Db 47 ACCTTCTCGAAGTCGTCTGCTCGGCTCGTCCAGGCTCG 6

RESULT 10  
US-09-385-707-1/c  
Sequence 1, Application US/09385707  
Patent No. 6238662  
GENERAL INFORMATION:  
APPLICANT: SCOTT, Hamish S.  
APPLICANT: ANSON, Donald S.  
APPLICANT: ORSBORN, Annette M.  
APPLICANT: NELSON, Paul V.  
APPLICANT: CLEMENTS, Peter R.  
APPLICANT: MORRIS, Charles P.  
APPLICANT: HOPWOOD, John J.  
TITLE OF INVENTION: SYNTHETIC ALPHA-L-IDURONIDASE AND GENETIC  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: SCULLY SCOTT MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/385,707

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/494,104  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8978Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Oligonucleotide DNA  
US-09-385-707-1

Query Match 13.6%; Score 18; DB 3; Length 47;  
Best Local Similarity 64.3%; Pred. No. 1e+04;  
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 61 ACCTTCTCGCGGCTGCTACTCGGCTGGATGAGCGCTCG 102  
Db 47 ACCTTCTCGAAGTCGTCTGCTCGGCTCGTCCAGGCTCG 6

RESULT 11  
US-09-639-696C-1/c  
Sequence 1, Application US/09639696C  
Patent No. 6524835  
GENERAL INFORMATION:  
APPLICANT: SCOTT, Hamish S.  
APPLICANT: ANSON, Donald S.  
APPLICANT: ORSBORN, Annette M.  
APPLICANT: NELSON, Paul V.  
APPLICANT: CLEMENTS, Peter R.  
APPLICANT: MORRIS, Charles P.  
APPLICANT: HOPWOOD, John J.  
TITLE OF INVENTION: SYNTHETIC ALPHA-L-IDURONIDASE AND GENETIC  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESS: NIXON PEABODY LLP  
STREET: 990 Stewart Avenue  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/639,696C  
FILING DATE: 16-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/191,171  
FILING DATE: 13-NOV-1998  
APPLICATION NUMBER: US 08/494,104  
FILING DATE: 23-JUN-1995  
APPLICATION NUMBER: AU PK9490/91  
FILING DATE: 14-NOV-1991  
APPLICATION NUMBER: AU PCT/AU92/00611  
FILING DATE: 12-NOV-1992  
APPLICATION NUMBER: US 08/084,254  
FILING DATE: 07-JUL-1993

```
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Pokalsky, Ann R.
;   REGISTRATION NUMBER: 34,697
;   REFERENCE/DOCKET NUMBER: 2249/304
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 516-832-7572
;   TELEFAX: 516-832-7555
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 47 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: Oligonucleotide DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-639-696C-1

Query Match      13.6%; Score 18; DB 4; Length 47;
Best Local Similarity 64.3%; Pred. No. 1e+04;
Matches 27; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 61 ACGTTCGCGGGCTGTCTACTCGGCTGGATGAGCGCTCG 102
      |||||
Db 47 ACGTTCGAAGTCGTCTGCTCGGCTGGTCCAGGTCCTCG 6

RESULT 12
US-09-581-105-9/c
; Sequence 9, Application US/09581105
; Patent No. 6603062
; GENERAL INFORMATION:
; APPLICANT: Schmidt, et al.
; TITLE OF INVENTION: Method For Producing Transgenic Plants With Modified 5-Aminolevulinic Acid Biosynthesis, And Method For Identifying 5-Aminolevulinic Acid Biosynthesis
; TITLE OF INVENTION: Acid Biosynthesis, And Method For Identifying 5-Aminolevulinic Acid Biosynthesis
; TITLE OF INVENTION: Acid Biosynthesis, And Method For Identifying 5-Aminolevulinic Acid Biosynthesis
; FILE REFERENCE: 514413-3827
; CURRENT APPLICATION NUMBER: US/09/581.105
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: PCT/EP98/08028
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: GERMANY 197 54 929.2
; PRIOR FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: primer ALAS1
US-09-581-105-9

Query Match      13.2%; Score 17.4; DB 4; Length 32;
Best Local Similarity 77.8%; Pred. No. 1.4e+04;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 35 GCGTGAGTTTGCAGCCCTGCATGCACA 61
      |||||
Db 30 GTGCCAGATTGTAGTCTCTGCATGCACA 4

RESULT 13
5240847-33
; Patent No. 5240847
; APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;
; ZOPHEL, ANDREAS; KRYSTEK, EDLTRAUD; MAURER-FOGY, INGRID;
; WICHE-CASTANON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTWANN, RUDOLF
; TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE
; (HMN-SOD)
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/167,261
; FILING DATE: 11-MAR-1988
```

```
;SEQ ID NO:33:
;LENGTH: 40
5240847-33

Query Match      13.2%; Score 17.4; DB 6; Length 40;
Best Local Similarity 68.6%; Pred. No. 1.4e+04;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 ATGCATCATCTGCCAGGACCTGCTGCAGGACTGCG 37
      |||||
Db 1 AAGCACTCTTTGCCAAGACTTGCCATACGACTACG 35

RESULT 14
PCT-US95-13975-35/c
; Sequence 35, Application PC/TUS9513975
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: McDonnell, Michael W.
; TITLE OF INVENTION: Recombinant Feline Herpes virus
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM 330 466 DX2
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13975
; FILING DATE: 26-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/329,883
; FILING DATE: 26-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39118-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0525
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human cytomegalovirus; Feline leukemia virus
PCT-US95-13975-35

Query Match      13.0%; Score 17.2; DB 5; Length 46;
Best Local Similarity 65.8%; Pred. No. 1.7e+04;
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 82 TCGGGCTGGATGAGCGCTCGTCCCTGTCTCTACTG 119
      |||||
Db 41 TCGGGCCCATTTGATCCATCGTCCCGGTGTCTCTATG 4

RESULT 15
5196308-1/c
; Patent No. 5196308
```

;  
; APPLICANT: NEPOM, GERALD T.; NEPOM, BARBARA S.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING THE DQW3.2  
; ALLELE ASSOCIATED WITH INCREASED RISK OF INSULIN-DEPENDENT  
; DIABETES MELLITUS  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/325,058  
; FILING DATE: 13-MAR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 745,321  
; FILING DATE: 14-JUN-1985  
; SEQ ID NO: 1:  
; LENGTH: 46  
5196308-1

Query Match 13.0%; Score 17.2; DB 6; Length 46;  
Best Local Similarity 73.3%; Pred. No. 1.7e+04;  
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 87 CTGGATGGAGCGCTCGTCCCTGTGTCTTAC 116  
Db 30 CTGGATGGAGATGGTCACTGTGGGTCC 1

Search completed: May 26, 2004, 09:20:41  
Job time : 26.4982 secs



```

Db      5 AGGTCCTGCAGCAGTATGGYACCAAGTTTGAAGCCTCCATCCA 47

RESULT 2
US-10-097-044A-18/c
; Sequence 18, Application US/10097044A
; Publication No. US20030143220A1
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/097,044A
; FILING DATE: 28-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0444P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8228
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-097-044A-18

Query Match      14.5%; Score 19.2; DB 15; Length 47;
Best Local Similarity 67.5%; Pred. No. 1.2e-04;
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      74 CTTGCTACTCGGCTGGATGGAGCGCTGTCCTGTCCTGCC 113
      |||||
Db      46 CTTGCTTCTGGTGTGCAACTGGCGCTCTCTCCACGACC 7

RESULT 3
US-10-097-044A-15
; Sequence 15, Application US/10097044A
; Publication No. US20030143220A1
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.

```

```

; Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/097,044A
; FILING DATE: 28-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0444P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8228
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-097-044A-15

Query Match      14.4%; Score 19; DB 15; Length 47;
Best Local Similarity 71.4%; Pred. No. 1.4e-04;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      74 CTTGCTACTCGGCTGGATGGAGCGCTGTCCTGCCCTG 108
      |||||
Db      12 CTTGCTTCTGGTGTGCAACTGGCGCTCTCTCCACG 46

RESULT 4
US-10-371-404-20
; Sequence 20, Application US/10371404
; Publication No. US20030194807A1
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: Walk-Through Mutagenesis
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/371,404  
FILING DATE: 20-Feb-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,623  
FILING DATE: 30-May-1995  
APPLICATION NUMBER: US 07/930,600  
FILING DATE: 05-APR-1991  
APPLICATION NUMBER: PCT/US91/02362  
FILING DATE: 05-APR-1991  
APPLICATION NUMBER: US 07/505,314  
FILING DATE: 05-APR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: RC90-01AY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 48 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-10-371-404-20

Query Match 13.8%; Score 18.2; DB 15; Length 48;  
Best Local Similarity 48.9%; Pred. No. 2.5e+04;  
Matches 23; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

QY 64 TTCTGCGGGTCTACTCGGCTGGTGGAGCGCTCGTCCCTGTG 110  
DB 1 TACTGCGGGGTARCTNCTTRGCGAGCASTTSGTMCYCKMCKYTTG 47

RESULT 5

US-10-131-827-1070  
Sequence 1070, Application US/10131827  
Publication No. US20040009479A1  
GENERAL INFORMATION:  
APPLICANT: Wohlgemuth, Jay  
APPLICANT: Fry, Kirk  
APPLICANT: Woodward, Robert  
APPLICANT: Ly, Ngoc  
TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES

FILE REFERENCE: 50661200120  
CURRENT APPLICATION NUMBER: US/10/131,827  
CURRENT FILING DATE: 2002-09-06  
PRIOR APPLICATION NUMBER: US 10/006,290  
PRIOR FILING DATE: 2001-10-22  
PRIOR APPLICATION NUMBER: US 60/296,764  
PRIOR FILING DATE: 2001-06-08  
NUMBER OF SEQ ID NOS: 9090  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1070  
LENGTH: 50  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-131-827-1070

Query Match 13.6%; Score 18; DB 16; Length 50;  
Best Local Similarity 70.6%; Pred. No. 2.9e+04;  
Matches 24; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 51 CTGCATGCACACGTTCTCGCGGCTTGCTACTCG 84  
DB 10 CTGTGTGAAGTCTTCTGCACTGCTTGCTAATGG 43

RESULT 6

US-10-076-802-81/c  
Sequence 81, Application US/10076802  
Publication No. US20030162249A1  
GENERAL INFORMATION:  
APPLICANT: Gray, Jeff  
APPLICANT: Buechler, Joe  
APPLICANT: Veeramallu, Uday K  
TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION

FILE REFERENCE: 014907-004700US  
CURRENT APPLICATION NUMBER: US/10/076,802  
CURRENT FILING DATE: 2002-02-13  
NUMBER OF SEQ ID NOS: 113  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 81  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine 5'  
OTHER INFORMATION: Heavy Chain Specific PCR Primer 188  
US-10-076-802-81

Query Match 13.5%; Score 17.8; DB 15; Length 44;  
Best Local Similarity 75.9%; Pred. No. 3.4e+04;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ACCTGCTGCACGACTGCGTGAGTTTGCGAG 48  
DB 35 ACCAGCTGCACCTTCGCTTGAGCCTGCAG 7

RESULT 7

US-10-076-802-110/c  
Sequence 110, Application US/10076802  
Publication No. US20030162249A1  
GENERAL INFORMATION:  
APPLICANT: Gray, Jeff  
APPLICANT: Buechler, Joe  
APPLICANT: Veeramallu, Uday K  
TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION

FILE REFERENCE: 014907-004700US  
CURRENT APPLICATION NUMBER: US/10/076,802  
CURRENT FILING DATE: 2002-02-13  
NUMBER OF SEQ ID NOS: 113  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 110  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine 5'  
OTHER INFORMATION: Heavy Chain Specific PCR Primer 208  
US-10-076-802-110

Query Match 13.5%; Score 17.8; DB 15; Length 44;  
Best Local Similarity 75.9%; Pred. No. 3.4e+04;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ACCTGCTGCACGACTGCGTGAGTTTGCGAG 48  
DB 35 ACCAGCTGCACCTTCGCTTGAGCCTGCAG 7

RESULT 8

US-10-076-802-111/c  
Sequence 111, Application US/10076802  
Publication No. US20030162249A1  
GENERAL INFORMATION:  
APPLICANT: Gray, Jeff  
APPLICANT: Buechler, Joe  
APPLICANT: Veeramallu, Uday K  
TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION

FILE REFERENCE: 014907-004700US  
CURRENT APPLICATION NUMBER: US/10/076,802  
CURRENT FILING DATE: 2002-02-13  
NUMBER OF SEQ ID NOS: 113  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 111  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine 5'  
OTHER INFORMATION: Heavy Chain Specific PCR Primer 208  
US-10-076-802-111

Query Match 13.5%; Score 17.8; DB 15; Length 44;  
Best Local Similarity 75.9%; Pred. No. 3.4e+04;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ACCTGCTGCACGACTGCGTGAGTTTGCGAG 48  
DB 35 ACCAGCTGCACCTTCGCTTGAGCCTGCAG 7

RESULT 9

US-10-076-802-111/c  
Sequence 111, Application US/10076802  
Publication No. US20030162249A1  
GENERAL INFORMATION:  
APPLICANT: Gray, Jeff  
APPLICANT: Buechler, Joe  
APPLICANT: Veeramallu, Uday K  
TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION

FILE REFERENCE: 014907-004700US  
CURRENT APPLICATION NUMBER: US/10/076,802  
CURRENT FILING DATE: 2002-02-13  
NUMBER OF SEQ ID NOS: 113  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 111  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine 5'  
OTHER INFORMATION: Heavy Chain Specific PCR Primer 208  
US-10-076-802-111

Query Match 13.5%; Score 17.8; DB 15; Length 44;  
Best Local Similarity 75.9%; Pred. No. 3.4e+04;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ACCTGCTGCACGACTGCGTGAGTTTGCGAG 48  
DB 35 ACCAGCTGCACCTTCGCTTGAGCCTGCAG 7

RESULT 10

US-10-076-802-111/c  
Sequence 111, Application US/10076802  
Publication No. US20030162249A1  
GENERAL INFORMATION:  
APPLICANT: Gray, Jeff  
APPLICANT: Buechler, Joe  
APPLICANT: Veeramallu, Uday K  
TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION

FILE REFERENCE: 014907-004700US  
CURRENT APPLICATION NUMBER: US/10/076,802  
CURRENT FILING DATE: 2002-02-13  
NUMBER OF SEQ ID NOS: 113  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 111  
LENGTH: 44  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Murine 5'  
OTHER INFORMATION: Heavy Chain Specific PCR Primer 208  
US-10-076-802-111

```

/ GENERAL INFORMATION:
/ APPLICANT: Gray, Jeff
/ APPLICANT: Buechler, Joe
/ APPLICANT: Veeramallu, Uday K
/ TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR PROKARYOTIC EXPRESSION
/ FILE REFERENCE: 014907-004700US
/ CURRENT APPLICATION NUMBER: US/10/076,802
/ CURRENT FILING DATE: 2002-02-13
/ NUMBER OF SEQ ID NOS: 113
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 111
/ LENGTH: 44
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
/ OTHER INFORMATION: Heavy Chain Specific PCR Primer 526
US-10-076-802-111

Query Match 13.5%; Score 17.8; DB 15; Length 44;
Best Local Similarity 75.9%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ACTGCTGCACGACTCGCTGAGTTTGCAG 48
||| ||||| ||||| |||||
Db 35 ACCACCTGCACGCTCCGCTTGAGCCTGCAG 7

RESULT 9
US-10-367-169-81/c
/ Sequence 81, Application US/10367169
/ Publication No. US20030228660A1
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Jeff
/ APPLICANT: Buechler, Joe
/ APPLICANT: Veeramallu, Uday Kumar
/ TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR POLYPEPTIDE EXPRESSION AND POLYPEPTIDE EXPRESSION
/ FILE REFERENCE: 11055US02
/ CURRENT APPLICATION NUMBER: US/10/367,169
/ CURRENT FILING DATE: 2003-02-13
/ PRIOR APPLICATION NUMBER: 10/076,802
/ PRIOR FILING DATE: 2002-02-13
/ NUMBER OF SEQ ID NOS: 113
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 81
/ LENGTH: 44
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
/ OTHER INFORMATION: Heavy Chain Specific PCR Primer 188
US-10-367-169-81

Query Match 13.5%; Score 17.8; DB 16; Length 44;
Best Local Similarity 75.9%; Pred. No. 3.4e+04;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 20 ACTGCTGCACGACTCGCTGAGTTTGCAG 48
||| ||||| ||||| |||||
Db 35 ACCAGCTGCACCTCCGCTTGAGCCTGCAG 7

RESULT 10
US-10-367-169-110/c
/ Sequence 110, Application US/10367169
/ Publication No. US20030228660A1
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Jeff
/ APPLICANT: Buechler, Joe
/ APPLICANT: Veeramallu, Uday Kumar
/ TITLE OF INVENTION: EUKARYOTIC SIGNAL SEQUENCES FOR POLYPEPTIDE EXPRESSION AND POLYPEPTIDE EXPRESSION
/ FILE REFERENCE: 11055US02
/ CURRENT APPLICATION NUMBER: US/10/367,169
/ CURRENT FILING DATE: 2003-02-13
/ PRIOR APPLICATION NUMBER: 10/076,802
/ PRIOR FILING DATE: 2002-02-13
/ NUMBER OF SEQ ID NOS: 113
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 111
/ LENGTH: 44
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Murine 5'
/ OTHER INFORMATION: Heavy Chain Specific PCR Primer 188
US-10-367-169-110

```



; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: PCT/EP98/08028  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: GERMANY 197 54 929.2  
; PRIOR FILING DATE: 1997-12-10  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 32  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: primer ALASI  
US-10-458-568-9

Query Match 13.2%; Score 17.4; DB 13; Length 32;  
Best Local Similarity 77.8%; Pred. No. 4.4e+04;  
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 35 CGGTGAGTTTCAGCCCTGCGATGCACA 61  
DB 30 GTGCCAGATTGTAGTCTCTGCATGCACA 4

RESULT 13  
US-10-216-054A-14  
; Sequence 14, Application US/10216054A  
; Publication No. US20030144232A1  
; GENERAL INFORMATION:  
; APPLICANT: CANCER RESEARCH VENTURES LIMITED  
; APPLICANT: AGAMI, Reuven  
; APPLICANT: BRUMMELKAMP, Thijs  
; TITLE OF INVENTION: EXPRESSION SYSTEM  
; FILE REFERENCE: KILBURN1100-1  
; CURRENT APPLICATION NUMBER: US/10/216,054A  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 60/377,482  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: UK 0130955.8  
; PRIOR FILING DATE: 2001-12-24  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 49  
; TYPE: RNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Predicted stem loop transcript  
US-10-216-054A-14

Query Match 13.2%; Score 17.4; DB 15; Length 49;  
Best Local Similarity 57.1%; Pred. No. 4.6e+04;  
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 83 CGGCTGATGAGCGCTCGTCCCTGTGTCCTACC 117  
DB 12 CGGGCCGAUCAAAGAGAGUCGGCCGAGUCCUGCC 46

RESULT 14  
US-10-324-184-14  
; Sequence 14, Application US/10324184  
; Publication No. US20030144239A1  
; GENERAL INFORMATION:  
; APPLICANT: CANCER RESEARCH VENTURES LIMITED  
; APPLICANT: AGAMI, Reuven  
; APPLICANT: BRUMMELKAMP, Thijs  
; TITLE OF INVENTION: EXPRESSION SYSTEM  
; FILE REFERENCE: KILBURN1100-2  
; CURRENT APPLICATION NUMBER: US/10/324,184  
; CURRENT FILING DATE: 2002-12-19  
; PRIOR APPLICATION NUMBER: US 10/216,054  
; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/377,482  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: UK 0130955.8  
; PRIOR FILING DATE: 2001-12-24  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 49  
; TYPE: RNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Sequence of the predicted stem loop transcript generated from  
; OTHER INFORMATION: the pSUPER-CDC20 vector as depicted in Figure 4.  
US-10-324-184-14

Query Match 13.2%; Score 17.4; DB 15; Length 49;  
Best Local Similarity 57.1%; Pred. No. 4.6e+04;  
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
QY 83 CGGCTGATGAGCGCTCGTCCCTGTGTCCTACC 117  
DB 12 CGGGCCGAUCAAAGAGAGUCGGCCGAGUCCUGCC 46

RESULT 15  
US-10-179-826-3  
; Sequence 3, Application US/10179826  
; Publication No. US20030235827A1  
; GENERAL INFORMATION:  
; APPLICANT: McKeown, Brian  
; TITLE OF INVENTION: Methods and Compositions for Monitoring Primer Extension and Polyn  
; TITLE OF INVENTION: Detection Reactions  
; FILE REFERENCE: 13164US  
; CURRENT APPLICATION NUMBER: US/10/179,826  
; CURRENT FILING DATE: 2002-06-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 22  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Primer - organism matches to Ovis aries  
US-10-179-826-3

Query Match 13.0%; Score 17.2; DB 16; Length 22;  
Best Local Similarity 86.4%; Pred. No. 5e+04;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 TCATCTGCCAGGACCTGTGCGCA 29  
DB 1 TCATGTGGCAGGAGCTGTGCA 22

Search completed: May 26, 2004, 10:22:40  
Job time : 577.198 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 05:12:53 ; Search time 916.382 Seconds

(without alignments)  
4301.492 Million cell updates/sec

Title: US-10-048-046-1\_COPY\_997\_1128

Perfect score: 132

Sequence: 1 acatgcacatcgcgcga.....ctaccgcgcgtccgcgtg 132

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 138346

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_mam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rod:\*\*

26: em\_gss\_phg:\*\*

27: em\_gss\_vrl:\*\*

28: gb\_gss1:\*\*

29: gb\_gss2:\*\*

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	19.2	14.5	49	9	AI284473
2	18.8	14.2	42	28	AZ788565
3	18.6	14.1	50	9	AU105739
4	18.6	14.1	50	9	AU105740

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	18.6	14.1	50	9	AU105741
6	18.2	13.8	49	9	AA886434
c 7	18.2	13.8	50	9	AU105564
8	17.8	13.5	50	28	AZ432474
c 10	17.6	13.3	49	9	AI941380
11	17.4	13.2	44	29	BX121674
12	17.4	13.2	48	28	AZ485794
c 13	17.4	13.2	50	28	BH624322
c 14	17.2	13.0	46	14	H45385
c 15	17.2	13.0	49	29	CG805551
c 16	17.2	13.0	50	9	AU105249
c 17	17.2	13.0	50	9	AU107541
c 18	17.2	13.0	50	28	AZ432474
c 19	17.2	12.9	50	9	AA948106
c 20	17.2	12.9	50	9	AU103176
c 21	17.2	12.9	50	9	AU105454
c 22	17.2	12.9	50	9	AU105455
c 23	16.8	12.7	50	9	AL788155
c 24	16.8	12.7	50	9	AU102427
c 25	16.6	12.6	43	9	AI744377
26	16.6	12.6	47	28	AZ795054
27	16.6	12.6	49	9	AA813480
28	16.6	12.6	50	9	AU105247
29	16.6	12.6	50	9	AU105248
30	16.6	12.6	50	9	AU105258
c 31	16.6	12.6	50	29	CG868958
c 32	16.4	12.4	35	29	DR20D78
c 33	16.4	12.4	37	14	H46233
c 34	16.4	12.4	42	9	AA878619
c 35	16.4	12.4	46	10	BE729845
c 36	16.4	12.4	46	28	BX171949
37	16.4	12.4	49	28	BH852307
38	16.4	12.4	49	28	BH852314
c 39	16.4	12.4	50	9	AU105060
c 40	16.2	12.3	42	29	DMES45174
41	16.2	12.3	43	12	BG707367
42	16.2	12.3	46	12	BG177798
43	16.2	12.3	48	28	AZ311191
c 44	16.2	12.3	50	9	AU102902
c 45	16.2	12.3	50	9	AU105570

#### ALIGNMENTS

RESULT 1  
AI284473/c  
LOCUS  
DEFINITION  
AI284473.1 49 bp mRNA linear EST 23-NOV-1998  
qu82a12.x1 NCI CGAP Eso2 Homo sapiens cDNA clone IMAGE:197858 3'  
similar to SW:FRPB HUMAN P02814 PROLINE-RICH PEPTIDE P-B ; contains  
element MER22 repetitive element ; mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLES  
JOURNAL  
COMMENT

AI284473  
AI284473.1 GI:3922706  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 49)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Nan Hu, M.D., Ph.D., Mark Roth, M.D., Phillip  
Taylor, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.lnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40UP from Gibco

High quality sequence stop: 1.

#### FEATURES

source  
1. .49  
/organization="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IWAGE:1978558"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP ESO2"  
/note="Organ: esophagus; Vector: pCMV-SPORT6; Site 1:  
Sal1; Site 2: NotI; Cloned unidirectionally. Primer:  
Oligo dr. Average insert size 1.1 Kb. Life Technologies  
catalog #: 11502-010"

#### ORIGIN

Query Match 14.5%; Score 19.2; DB 9; Length 49;  
Best Local Similarity 62.5%; Pred. No. 5e+05;  
Matches 30; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 84 GGGCTGATGAGCGCTCGTCCCTGTCTACTGCGCTGTCCCGT 131

Db 48 GGGCGGGGGGGATTCTTCCCGCGCCCTACTTCCCGCCGACCCCTT 1

#### RESULT 2

AZ788565

LOCUS AZ788565 42 bp DNA linear GSS 16-FEB-2001  
DEFINITION 2M0035F14R Mouse 10kb plasmid UUGCIM library Mus musculus genomic  
c1one UUGC2M0035F14 R, genomic survey sequence.

ACCESSION AZ788565

VERSION AZ788565.1 GI:12928495

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 42)

REFERENCE

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D. Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

TITLE

JOURNAL

COMMENT

Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0035 row: F column: 14  
Seq primer: CACACGGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 42.

#### FEATURES

source  
1. .42  
/organization="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0035F14"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGCIM library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 [gi|4732114|gb|AF129072.1], a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

#### ORIGIN

Query Match 14.2%; Score 18.8; DB 28; Length 42;  
Best Local Similarity 76.7%; Pred. No. 5.9e+05;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 39 GAGTTTCAGCGCTGCATGCACACGTTCTG 68

Db 7 GTGTTGAATCCCTGCATCTACAGTGATG 36

#### RESULT 3

AU105739

LOCUS AU105739 50 bp mRNA linear EST 30-AUG-2001  
DEFINITION AU105739 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone  
HRC01265, mRNA sequence.

ACCESSION AU105739

VERSION AU105739.1 GI:13555260

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 50)

REFERENCE

AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J.,  
Hata, H., Ota, T., Isegai, T., Tanaka, T., Morishita, S., Okubo, K.,  
Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.  
Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp  
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and  
Sugano, S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),  
149-156 (1997).

#### FEATURES

source  
1. .50  
/organization="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRC01265"  
/clone\_lib="Sugano Homo sapiens cDNA library"

#### ORIGIN

Query Match 14.1%; Score 18.6; DB 9; Length 50;  
Best Local Similarity 84.0%; Pred. No. 7.1e+05;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 107 TGTGTCTTACCTGCGCTGTCCCGT 131

```

Db      10 TGTGTCCTGTGACGCGCTGTCCCT 34
|||||
RESULT 4
AUI05740
LOCUS
DEFINITION AUI05740 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HS105409, mRNA sequence.
ACCESSION AUI05740
VERSION AUI05740.1 GI:13555261
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
Sugano,S. Construction and characterization of a full
length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
149-156 (1997).

FEATURES
Location/Qualifiers
source 1..50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match 14.1%; Score 18.6; DB 9; Length 50;
Best Local Similarity 84.0%; Pred. No. 7.1e+05;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 107 TGTGTCCTACCTGCGCTGTCCCGT 131
|||||
Db 11 TGTGTCCTGTGACGCGCTGTCCCT 35
|||||

RESULT 6
AA886434
LOCUS
DEFINITION AA886434 49 bp mRNA linear EST 30-MAR-1998
Oj52d09.s1 NCI CGAP Kid3 Homo sapiens cDNA clone IMAGE:1501937 3'
similar to TR:Q04118 Q04118 SALLIVARY PROLINE-RICH GLYCOPROTEIN G1
PRECUSOR. ;contains MER22.t3 MER22 repetitive element ;, mRNA
sequence.
ACCESSION AA886434
VERSION AA886434.1 GI:3001542
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES
Location/Qualifiers
source 1..49
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1501937"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid3"
/notes="Organ: kidney; Vector: p77T3D-pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not

```

```

COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
Sugano,S. Construction and characterization of a full
length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
149-156 (1997).

FEATURES
Location/Qualifiers
source 1..50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match 14.1%; Score 18.6; DB 9; Length 50;
Best Local Similarity 84.0%; Pred. No. 7.1e+05;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 107 TGTGTCCTACCTGCGCTGTCCCGT 131
|||||
Db 11 TGTGTCCTGTGACGCGCTGTCCCT 35
|||||

RESULT 6
AA886434
LOCUS
DEFINITION AA886434 49 bp mRNA linear EST 30-MAR-1998
Oj52d09.s1 NCI CGAP Kid3 Homo sapiens cDNA clone IMAGE:1501937 3'
similar to TR:Q04118 Q04118 SALLIVARY PROLINE-RICH GLYCOPROTEIN G1
PRECUSOR. ;contains MER22.t3 MER22 repetitive element ;, mRNA
sequence.
ACCESSION AA886434
VERSION AA886434.1 GI:3001542
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES
Location/Qualifiers
source 1..49
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1501937"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid3"
/notes="Organ: kidney; Vector: p77T3D-pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not

```



```

ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 49)
REFERENCE
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Parson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand This clone is
available through: Resgen, Invitrogen Corp. 2130 South Memorial
Parkway Huntsville, AL 35801 For further information call:
(800)-533-4363 or contact via email: ccu@resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..49
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl013-453"
/tissue_type="Whole seedlings, 2-3 week old seedlings,
greenhouse grown"
/lab_host="XL10-Gold"
/clone_lib="Gm-cl013"
/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from whole seedlings of 2-3 week old greenhouse grown
plants. The cDNA library was prepared using the Stratagene
pBluescript II XR cDNA library construction kit.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly (dT) sequence with a XhoI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into XL10-Gold host cells. This library
was constructed by Dr. Randy Shoemaker and Dr. John
Erpelding."
ORIGIN
Query Match 13.3%; Score 17.6; DB 9; Length 49;
Best Local Similarity 65.0%; Pred. No. 1.2e+06;
Matches 26; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 40 AGTTTCAGCGCTGCATGCACACGTTCTCGCGGCTTGCT 79
DB 10 ATTGTGTTCCCTGGCAGCTCACCTCTCGTGGCTTGAT 49
RESULT 10
BX121674/c 44 bp DNA linear GSS 28-JAN-2003
LOCUS Danio rerio genomic clone DKEY-61M14, genomic survey sequence.
DEFINITION BX121674
ACCESSION BX121674.1 GI:27952599
VERSION GSS.
KEYWORDS Danio rerio (zebrafish)
SOURCE Danio rerio
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 44)
AUTHORS Humphray,S.J., Huckle,E. and Durham,J.L.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 61M14. 61M14 is
part of the DanioKey BAC Library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/projects/D_rerio/.
FEATURES
Location/Qualifiers
1..44
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-61M14"
/tissue_type="Testis"
/notes="vector pindigoBAC-536"
ORIGIN
Query Match 13.2%; Score 17.4; DB 29; Length 44;
Best Local Similarity 61.8%; Pred. No. 1.3e+06;
Matches 21; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 76 TGCTACTCGGCTGGATGGAGCGCTGTCCTCT 109
DB 41 TGCTGCTCGGNGGTGNGTGNNGTCCTCT 8
RESULT 11
AZ485794 48 bp DNA linear GSS 05-OCT-2000
LOCUS IM0313K24F Mouse 10kb plasmid UGCLM library Mus musculus genomic
DEFINITION clone UGCLM0313K24 F, genomic survey sequence.
ACCESSION AZ485794
VERSION AZ485794.1 GI:10651939
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 48)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Rielly,M., Rose,M., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright D. Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: cdunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0313 row: K column: 24
Seq primer: CGTTGTAACACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 48.
FEATURES
Location/Qualifiers
1..48
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCLM0313K24"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

```

/clone lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## ORIGIN

Query Match 13.2%; Score 17.4; DB 28; Length 48;  
 Best Local Similarity 62.8%; Pred. No. 1.4e+06;  
 Matches 27; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 49 CCTGATGCACAGCTTCGCCGCTTCTACTCGGCTGGA 91  
 |||||  
 Db 2 CCTGGTCCCATGATCGCGGGCGCTCGCTGAGTTGGA 44

## RESULT 12

BH624322  
 LOCUS BH624322 50 bp DNA linear GSS 30-JAN-2002  
 DEFINITION 1007078A02.2EL\_Y2 1007 - RescueMu Grid H Zea mays genomic, genomic survey sequence.

## ACCESSION

VERSION BH624322

## KEYWORDS

SOURCE GSS.

## ORGANISM

Zea mays  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 50)

## REFERENCE

Walbot V.  
 Maize genomic sequences found using engineered RescueMu transposon Unpublished (2001)

## AUTHORS

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1007078 row: 30

Class: transposon-tagged.

## FEATURES

source

1..50  
 Location/Qualifiers

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A188/B73"

/db\_xref="taxon:4577"

/tissue type="leaf"

/dev stage="adult"

/lab\_host="DH10B"

/clone\_lib="1007 - RescueMu Grid H"

/note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA."

Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

## ORIGIN

Query Match 13.2%; Score 17.4; DB 28; Length 50;  
 Best Local Similarity 62.8%; Pred. No. 1.4e+06;  
 Matches 27; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Oy 89 GGATGGAGCGCTGCTCCTGCTCTACCTCGCGCTGTCCTCGT 131  
 |||||  
 Db 1 GGCTGGAGCAGTGTCTTTGGCCGGATTGTCTCTCTTGT 43

## RESULT 13

H45385/c

## LOCUS

H45385 46 bp mRNA linear EST 31-JUL-1995  
 YN9912.r1 Soares adult brain N2b5HB53r Homo sapiens cDNA clone IMAGE:176566 5' similar to SP:TFS2\_MOUSE P10712 TRANSCRIPTION FACTOR S-II ; mRNA sequence.

## ACCESSION

VERSION H45385.1

KEYWORDS GI:921437

SOURCE EST.

ORGANISM Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 46)

REFERENCE Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

AUTHORS

## TITLE

The WashU-Merck EST Project

Unpublished (1995)

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Seq primer: M13RP1

High quality sequence stop: 1.

Location/Qualifiers

1..46

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:3838762"

/db\_xref="taxon:9606"

/clone="IMAGE:176566"

/sex="Male"

/dev stage="55-year old"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares adult brain N2b5HB53r"

/note="Organ: brain; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTACCACTGTAAGTCGAGCGCGCCGCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector

(Pharmacol). Library went through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M.Fatima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."

## ORIGIN

Query Match 13.0%; Score 17.2; DB 14; Length 46;  
Best Local Similarity 65.8%; Pred. No. 1.5e+06;  
Matches 25; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 5 GCATCACTGCCAGGCTCTCCAGCAGCTGCGTGAGT 42  
Db 44 GGAGCTTCTTCCAGGACTTGATGAGAGACTTGCCAGT 7

## RESULT 14

CG805551/c  
LOCUS 1118060E02.2EL\_y1 1118 - RescueMu Grid S Zea mays genomic, genomic survey sequence.  
DEFINITION 49 bp DNA linear GSS 10-NOV-2003

ACCESSION CG805551

VERSION CG805551.1 GI:38241970

KEYWORDS GSS

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 49)

AUTHORS Walbot,V.

TITLE Maize genomic sequences found using engineered RescueMu transposon

JOURNAL Unpublished (2001)

COMMENT Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 1118060 row: 36

Class: transposon-tagged.

Location/Qualifiers

source

1. 49

/organism="Zea mays"

/mol\_type="genomic DNA"

/cultivar="mixed background W23/A188/B73"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

/dev\_stage="adult"

/lab\_host="DH103"

/clone\_lib="1118 - RescueMu Grid S"

/notes="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site: 1: BamHI; Site: 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu'.

'RescueMu', Grid S was grown at San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH103 cells were transformed and then screened on LB plates with ampicillin."

## ORIGIN

Query Match 13.0%; Score 17.2; DB 29; Length 49;  
Best Local Similarity 73.3%; Pred. No. 1.5e+06;  
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 23 TGTGTACGACTGCGTGAGTTTGCAGCCCT 52

Db 34 TGCAGCATGAGTGTGTGATATGCGCAGCT 5

## RESULT 15

LOCUS AU105249

DEFINITION AU105249 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC05540, mRNA sequence.  
ACCESSION AU105249

VERSION AU105249.1 GI:13554770

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Kato,H., Oca,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.

TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites

JOURNAL EMBO Rep. 2 (5), 388-393 (2001)

MEDLINE 21270072

PubMed 11375929

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S.

Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

Location/Qualifiers

source

1. 50

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="HRC05540"

/clone\_lib="Sugano Homo sapiens cDNA library"

ORIGIN

Query Match 13.0%; Score 17.2; DB 9; Length 50;

Best Local Similarity 60.9%; Pred. No. 1.6e+06;

Matches 28; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 40 AGTTTCAGCCCTGATGACACAGCTTCGCGCGGCTTGCTACTCGG 85

Db 1 AGTTTCAGCCCTGATGACACAGCTTCGCGCGGCTTGCTACTCGG 46

Search completed: May 26, 2004, 09:17:48

Job time : 919.382 secs